CONTROL CHARTS

IN FACTORY MANAGEMENT

WILEY MATHEMATICAL STATISTICS SERIES

Walter A. Shewhart, Editor

Applied Statistics

COCHRAN and COX—Experimental Designs.

DODGE and ROMIG—Sampling Inspection Tables.

RICE—Control Charts.

Mathematical Statistics

HOEL—Introduction to Mathematical Statistics.

WALD—Sequential Analysis.

Related Books of Interest to Statisticians

HAUSER and LEONARD—Government Statistics for Business Use.

CONTROL CHARTS

IN FACTORY MANAGEMENT

By

WILLIAM B. RICE

Consulting Business Statistician

NEW YORK

JOHN WILEY & SONS, INC.

CHAPMAN & HALL, LTD.

LONDON

Second Printing, April, 1950

PRINTED IN THE UNITED STATES OF AMERICA

To

W. EDWARDS DEMING

Who introduced the author to statistical
quality control and who has given him
continued encouragement in his work

PREFACE

Control charts, by their simplicity and fact-finding power, have amply proved their value as one of the most useful tools of factory management. The author has tried to demonstrate in these pages the basic function of statistical quality control in manufacturing plants: to help remove causes of bad work by studying and analyzing factory processes. How the control-chart technique can help to build economical quality into a product is the theme of this book.

During recent years the use of control charts spread so rapidly that the need has arisen for a book to bridge the gap between W. A. Shewhart's *Economic Control of Quality of Manufactured Product* and the American Standard Association's Z1.1, Z1.2, and Z1.3 pamphlets. The former is on an advanced mathematical plane, and the latter is on the practical level of an operating handbook. The present volume, *Control Charts in Factory Management*, is written with the hope that it will in some measure bridge that gap. Based as it is upon several years' working and teaching experience, the author believes that it will add something worth while to the all-too-meager literature on the subject.

To those who have never been exposed to statistical ideas this book offers a practical introduction to a fertile field. To those who are already familiar with control-chart techniques as presented in the ASA pamphlets and elsewhere, it offers enough theory to help them deepen their appreciation and to lead them on to more complex applications. To those who have the technical equipment for grasping the full implications of the Shewhart technique, this book may open up a wider, more complete view of the work-a-day potentials inherent in the control-chart method.

To the business executive, the man who carries the heavy burden of top-management responsibility, this book is particularly dedicated. If in reading it he grasps the power of control charts to ease his problems, it will have served its purpose. Executives need not understand the mathematics, nor even the methods of practical application. The top factory executive need only absorb the philosophy of control expressed in these pages and give full support to the engineer, production man, inspector, or statistician who is assigned the responsibility of translating ideas into action in his own plant. Toward this end the author has

striven to achieve both clarity and brevity, because above all he wishes to reach the busy practical men who run our factories. They are the ones who can make most effective use of statistical quality control.

Although the author has placed major emphasis upon practical applications, especially in the case histories, some statistical formulas have been unavoidable. There are several places in this book where the mathematical statistician will see something to criticize from the viewpoint of exact statement. It should be kept in mind that the author is trying to reach, through laymen's language, the man on the job who is not so much interested in foolproof statements as he is in getting a picture of how he can use control charts to advantage in his work.

The author is deeply indebted to Walter A. Shewhart for his editing of the manuscript.

WM. B. RICE

Pasadena
December, 1946

CONTENTS

INTRODUCTION

INTRODUCTION

STATISTICAL QUALITY CONTROL: WHAT IT IS AND WHAT IT DOES

The word *control,* when used in business, implies a procedure by which management can accumulate and interpret the facts necessary for determining whether or not business activities are proceeding according to plan. Particularly, control means a method of determining when the activities of business are deviating more than they should from the course laid down by management. In this sense, for instance, production control provides the tool whereby management can plan deliveries in advance. It sets up a procedure whereby management need not be concerned so long as production is moving along smoothly, but, if there is any significant deviation from the desired results, it is automatically called to management's attention in order that steps can be taken for its correction. Sales control, likewise, calls for a means by which sales can be measured against predetermined management-set standards. Quality control, by analogy, is a method of measuring the quality of a factory's output against standards that are determined by a consideration of all the factors—the requirements of the sales department, the capabilities of the processes, the skill of the workmen, the availability of raw materials, and the engineering design—that will produce a maximum profit for the business organization as a whole.

The word *quality* usually implies a high level of perfection in the finished product. A *quality line* means, in the ordinary sense, a product that has superior consumer appeal. This is not the sense in which the word quality is used here. *Quality* does not imply any particular degree of desirability; it is simply a name attached to a characteristic or combination of characteristics of a manufactured product. When the characteristics are compared with a standard—a standard set by management—the quality can be said to be good if it meets the standard, and to be poor if it is either better or worse than the standard. As an illustration of this definition, consider the problem that faced a manufacturer of a certain small hard-rubber product. This particular rubber item was made by the millions every month and was used for a purpose that allowed dimensions and surface defects a wide range of

variation, but called for a specified hardness. The manufacturer had two alternatives. He could use scrap raw material, allow a large percentage of the parts to fail, and sell the product at half price, or he could use new rubber, guarantee almost perfect performance as to hardness, and sell the product at full price. His profit margin in each case would be the same. In the first case, according to the accepted meaning of the word quality, the product would be poor; in the second case it would be good. However, according to the meaning of the word as it is used here, both qualities would be good, because both would meet the profit standard set by management. Which of the two manufacturing methods should be used would have to be determined by management on the basis of salability. That is, would it be more economical for purchasers of the product to buy the cheaper parts and use more of them, or to buy the more expensive parts and use fewer of them? If, in the first case, a standard of 25 per cent defective were allowed, and the price were 50 cents per thousand, it might pay the purchasers to buy the 25 per cent defective material rather than a practically perfect product that cost one dollar per thousand. It is problems such as these that make it necessary to define the word quality in a concrete and definite sense as a comparison of objective characteristics of the product with management-set standards for those characteristics. With this meaning the words *acceptable quality* apply to a product whose objective measurable characteristics conform to the standard. If the product is either superior or inferior to the standard, it is unacceptable, because in the one case it probably would be too expensive to produce, and in the other case it would not meet the purchasers' requirements.

Quality control in the broad sense requires the co-ordination of consumer research, product research, and production research, with special emphasis on cost, directed towards meeting the standards set by management. In the operation of quality control three distinct steps are involved. First, accurate, correct, and adequate facts should be gathered, and these facts then should be properly analyzed and interpreted. Second, from these facts and the conclusions drawn from them, management should set the standard of quality for the outgoing product. Third, control procedure should be set up in order to assure management that the operations of the business are being conducted at the desired level, and that any significant deviation therefrom is corrected promptly. *Statistical* quality control means the application of statistical methods as a scientific technique for collecting and analyzing the data, for setting the standards, and for maintaining ad-

herence to the standards. Statistical quality control is a method of applying statistical techniques to the collection and analyzing of inspection and other data in order to achieve and maintain maximum economy in manufacturing processes.

STATISTICAL TECHNIQUES

This concept of statistical quality control includes several significant ideas. First of all, it uses statistical techniques. Statistics is the science of collecting, measuring, analyzing, and interpreting quantitative facts. As applied to the quality characteristics of a manufactured product it makes possible the establishment of sound norms or standards, and enables management to differentiate clearly between deviations which (*a*) should be expected by chance in the operation, or (*b*) are too large to occur by chance alone. For instance, a salesman sold $5,500 worth of goods last month but only $3,500 this month; should he be complimented on last month's record, should he be discharged because of this month's record, or should nothing be said? From the two months' data given, it would be impossible to tell. If records of the salesman's monthly sales for several years back were available, it would be possible by statistical techniques to show whether or not this month's sales were smaller than normally should be expected, in view of the variation in his previous record. Similarly, when a machine produces a part that is too large or too small, according to specifications or other standards, it is possible, from an engineering study and statistical analysis of the operation, to determine whether or not that piece was one that the machine could normally be expected to produce. Or, suppose that a workman produces 5 per cent bad work today against only 1 per cent yesterday: it may be difficult to tell whether or not he should be reprimanded. If, for instance, during any month his daily record varies from 0 to 9 per cent bad work, common sense alone would indicate that he should not be blamed for a 5 per cent day. By drawing definite limits between what variation can be expected and what is unexpected, and by comparing these limits with the variations that have been determined by management to be economical, statistical techniques enable management to take action in the right direction when troubles are found to exist. The control chart, which is the subject of this book, is one of the simplest, most practical, least expensive, and most reliable techniques ever devised for helping Management to keep economic control of its process.

ANALYSIS OF FACTS

Statistical quality control requires the analysis of facts. In most manufacturing plants the collection of quantitative data about the product is a function of the inspection department. Statistical quality control, therefore, should start with inspection, for the purpose of making sure that the facts collected are adequate and correct. In this respect there are two phases of the statistical quality control function. First, the *routine* accumulation of inspection records should be so organized that the necessary facts can be analyzed quickly and easily. This may involve such changes in the inspection department as an evaluation of the various inspectors' abilities, the installation of sampling plans, the setting up of process inspection procedures, and the checking of the inspection department's work in order to ascertain whether or not inspection is providing data with required accuracy and reliability. Second, statistical quality control often requires that special inspections be undertaken in special ways in the course of special studies of factory operations. This is a particularly important function of quality control, because it is similar to experimental work in which the design of the experiment, that is, the way in which the data are collected, can be made most efficient by the application of modern statistical techniques to the planning of the experiment. In both routine and special phases of inspection, the purpose is primarily to deduce the nature of a given process from measurements and observations made on the product of that process. In other words, the character of the man is judged by his appearance. This is, in fact, one of the best ways in which the characteristics of a process can be judged. As an illustration, a certain soap manufacturer had his chemists devise a new formula with a supposedly special skin-protecting quality. He tested the first cakes of soap produced by the new formula and found that they actually did possess the properties desired. But if he had not tested the soap he would not have known with certainty whether the soap did have the characteristics he wanted. As another illustration, an open-end wrench was designed with certain specifications, among them the requirement that the sides of the jaws should be parallel. It was specified that the openings could be either broached or milled. It was found, however, that neither of these processes produced openings with parallel jaws. The broaching operation caused a toe-in at the points, and the milling operation caused a toe-out at the points. After considering the use for which the wrenches were designed, management came to the conclusion that neither the toe-in nor

the toe-out was large enough to affect the usefulness of the wrenches; therefore, both operations were permitted. No matter how good the designing is, nor how perfect the planning, the production has to be done by people who are fallible and by machines that are not perfect. Therefore it never can be assumed that the product is all that it is desired to be. The only way to find out in what respect and how much the manufactured product deviates from the designed product is by observation of the physical characteristics of the product. This is the inspection function.

QUALITY OF THE PROCESS

When statistical quality control is used in a plant, the first step, as previously pointed out, is to make sure that the data collected are adequate and of the right kind. The collection of data is merely a means to an end. The next step is to infer from the characteristics of the product what the characteristics of the process are. Here the real meaning of the word quality as used in statistical quality control becomes clear. The word quality refers much more to the quality of a *process* than it does to the quality of a product. If a process is capable of meeting the specifications set by management, and with the economy necessary for maintaining acceptable costs, then the quality of the process is good. An automatic screw machine will illustrate this point. If the tolerance on a certain automatic turning is ±0.001 inch, and a six-spindle automatic can maintain not better than ±0.002 inch, then the process is not satisfactory and the quality is not good. In this situation it may be necessary to change the design tolerances, or to use a different process. It is often possible, however, by means of control charts, to reduce the variability in a six-spindle automatic to the region of ±0.001 inch. In that case the process will be producing what management desires, and the quality of the process will be acceptable. On the other hand, suppose that the management-set tolerances are ±0.003 inch. If the machine can hold ±0.002 inch, then greater efficiency may be achieved by allowing the tools to run a little longer, or by using an older and less accurate machine. The importance of knowing what the operation *does*, what it *can do*, and what it *must do*, if maximum efficiency is to be obtained, is self-evident. The techniques of statistical quality control are well adapted to solving this type of problem, and to giving management the information it needs for making sound decisions regarding the maximum profit potential of its business.

OPERATING STANDARDS

When inspection data have been analyzed and applied to improving and controlling an operation, management is in a position to set its standards. Before standards can be set, however, the operation must have become so stabilized that management can be reasonably sure that the standard will be maintained. This requires that a state of *statistical control* be established. If control charts have been run on an automatic screw machine and they show that the machine is capable of producing a certain part within limits of ±0.001 inch, and, if the chart shows that these limits have been held for hours at a time, then management can predict confidently that future operations also will maintain those limits. Then, management safely can set a standard for that operation of ±0.001 inch, with the knowledge that a minimum of down time will be required. If, however, management does not know that the operation can maintain ±0.001 inch for a long period of time—that is, if parts frequently are made too large or too small—then management cannot confidently set a limit of ±0.001 inch, because the maintaining of these limits may require an excessive amount of machine adjustment. A state of statistical control also can be called a state of *valid predictability*, in which the *future* product quality of the operation can be predicted with a high degree of confidence from the characteristics of the product already produced.

When a state of statistical control has been achieved and management has been able to set economical, valid, and reliable standards, the procedure for making sure that the standards are maintained should be set up. This can be done through the techniques of statistical quality control, with less attention by management, than would be possible if statistical control were not achieved. Primarily, the same statistical techniques that are used to *achieve* control also can be used to maintain that control.

PROFIT CONTROL

Quality control perhaps can better be called *profit control.* Economy can be gained by improving the process, or by improving inspection methods, or by changing the design, or by changing the user's idea of what he wants. A few examples will illustrate the type of improvement that can be made in each of these ways.

1. Improving the process. A plant that specialized in automatic screw machine turnings had about 100 screw machines, some new, some old. It was difficult to schedule the work because often the machines were unable to hold the tolerances required. A great deal

of blame was laid on the old machines, and it was proposed to buy new ones in their stead. On one particularly difficult job, where the specifications called for a tolerance range of 0.002 inch, inspection showed that in a run of 1,000 pieces the difference in size between the largest and smallest was 0.006 inch. Since this particular job was done on one of the old machines, it had been decided to replace it with a new one, at the time when statistical quality control was introduced into the plant. One of the first studies undertaken was on this particular operation. Control charts were kept on the machine over a period of about three weeks, at the end of which time many minor maladjustments had been corrected and the operating range of the machine had been reduced to 0.002 inch. The result was so worth while that the same technique was applied to the other old machines, with similar results. It was soon discovered that the old machines now were producing work with a smaller range of size than the new machines. Studies then were conducted on the new machines, and similar improvements were made. On most of the machines it was found that the principal cause of the excessive variation was differences among the spindles. When these differences had been removed, most of the causes of trouble were eliminated. The control-chart technique, applied in this case, made it unnecessary for the company to invest large sums of money in new equipment; in fact, with their old equipment they were doing closer work than most of their competitors who had new machines.

The people who work at machines are certainly a part of the process. Control charts often have brought about improvements in the quality of work, where handwork is the major factor. In a certain plant, there were about 30 men working in the polishing department. Two of these men, Mr. A and Mr. B, vied with each other week after week for the worst quality record. When statistical-quality-control procedures were applied in the polishing department, the daily records of inspection were plotted on control charts, one chart for each workman. Mr. A and Mr. B both averaged about 9 per cent defective work. After sufficient time had elapsed for the charts to tell the story, the department head called these two men into his office to show them the charts. After this discussion, Mr. A showed an improvement overnight from an average of 9 per cent bad work down to 2 per cent; Mr. B showed practically no improvement. Shortly thereafter both men were working side by side for two days on the same job. Mr. A averaged about 2½ per cent bad work, and Mr. B about 12 per cent. The latter, who had a violent temper, flew into a rage when accused of not doing good work. He insisted that his work was just as good as Mr.

A's. When Mr. B was taken out to the polish wheel and asked whether a certain piece was up to standard, he said it was. The department head insisted that any good polisher would know that it was not up to standard. When the argument was over, Mr. B was advised to go to an oculist. There he discovered that his eyesight was poor, and that the reason he was doing bad work was that he could not tell the good work from the bad. A pair of glasses solved his problem and enabled him to keep pace with his competitor, Mr. A.

2. Improving inspection. In almost every plant statistical quality control can bring improvements and economies in inspection that could not be achieved in any other way. Often products are inspected 100 per cent that need only to be sampled; sometimes sampling is done on characteristics that should be detail-inspected. Sometimes the qualities that need real inspection are not even realized. A case of this sort arose in the assembly of a ball and spring into a hole that was drilled in a metal bar. The assembly process consisted of dropping the spring into the hole and of crimping the ball in on top of it. Trouble was being encountered because often the ball heights would run from 25 to 50 per cent higher than the tolerances allowed. A detailed and lengthy statistical analysis of the product was undertaken, including a study of the variability in the crimping process and correlation analysis of the height of the balls and the depth and diameter of the holes into which they were crimped. As a result of this investigation, it was discovered that the diameter of the hole was the controlling factor in the variability of ball height: the larger the hole, the higher the ball; the smaller the hole, the lower the ball. As a result of this finding, and in order to control the diameter of the holes as economically as possible, a scientific process-sampling procedure was set up by which the operator could be sure that the holes were maintained to tolerances with a minimum of time spent in gauging. Thereafter, practically no high balls appeared in the finished product.

In another case, a problem arose in connection with the straightening of metal bars. These bars were made in various sizes, some very large and some very small. It was general practice to straighten the large ones after they had been annealed, when they were soft, and then to inspect them 100 per cent after heat treatment and other operations, in order to pick out those that had been made crooked subsequent to annealing. With the small bars, however, it was general practice to send them straight through to heat treatment and final inspection and to straighten only those that were too crooked to meet the standard; these had to be straightened hard, which was a much more difficult and costly process. A high percentage of crooked bars had been found

at final inspection for many months. The problem arose, when should the bars be straightened soft, and when hard? A study was made, for each size of bar, of the unit cost of straightening when soft and straightening when hard. A simple straightness gauge also was designed. With the cost data at hand, and with the new gauge, which provided an objective quantitative measure of straightness, it was possible to determine for each size of bar what percentage could be allowed to pass on most economically to final inspection without being straightened when soft. In order to make sure that the straighteners did not undertake to straighten when it was not necessary, nor pass lots that needed straightening, a sample inspection was made on every lot just prior to the straightening operation. When the sample inspection showed that the percentage of crooked pieces was larger than the economical percentage allowed, that lot was sent in to the straighteners to be straightened when soft; otherwise, the lot was passed on to final inspection, and the crooked pieces were rejected there and sent back for straightening hard. During this study, control charts were kept on each of the straighteners and on the straightening press. Among other things, it was discovered that the press needed a complete overhaul, and that one of the straighteners—the one who had been employed longest—was doing a very inferior grade of work. When the faults were corrected, the cost of straightening was reduced by almost 50 per cent.

3. Improving the design. It is often difficult to decrease the variability or *scatter* in a manufacturing process. The newly employed manager of an aircraft-parts plant realized the truth of this statement shortly after he had assumed his new responsibilities. He discovered that far too large a proportion of production cost was being charged to reworking and repairing parts that had failed to pass inspection. A statistical investigation revealed process variations up to five times the tolerance ranges specified on the blueprints. An overhaul of engineering designs naturally was the result of these findings. In the case of a shaft-and-bearing assembly, the shaft had an engineering tolerance range of 0.0005 inch, and the bearing a tolerance range of 0.0008 inch. Analysis of the manufacturing process showed that, so far as production was concerned, the reverse situation was true: there was an operating tolerance range on the shaft of 0.0007 inch and on the bearing of 0.0004 inch. Because the design and production tolerances did not agree, rework costs had been excessive. Since it was easier to change the blueprints than to change the process in this case, the specifications simply were reversed. No harm was done to the assembly—minimum and maximum clearances between shaft and bearing were still the same—

but practically all rework on account of the shaft failing to gauge was eliminated. The accompanying table illustrates how this problem was solved to the benefit of all concerned.

	Before Correction		After Correction	
	Engineering Tol. Range (Inch)	Operating Tol. Range (Inch)	Engineering Tol. Range (Inch)	Operating Tol. Range (Inch)
Shaft	0.0005	0.0007	0.0008	0.0007
Bearing	0.0008	0.0004	0.0005	0.0004

Before the correction was made the shaft-manufacturing process had overflowed its specifications by 0.0002 inch, while the bearing process had too much room by 0.0004 inch. After correction both were inside their tolerances.

It is sometimes possible to open up engineering tolerances with benefit to all and harm to none. One instance of this sort occurred in connection with the assembly of a metal bar into a block. The square hole in the block was hot-punched, but the square that fitted the hole on the end of the rod was upset. The tolerance on both operations was ±0.002 inch. The inspection performed at the upsetting process resulted in frequent stoppages of the job because of parts failing to gauge. These stoppages became so frequent as to hinder seriously the work of that department. A comparative analysis of the inspection records at the upsetting and assembly operations revealed that no failures in assembly were discovered. The discrepancy in these two inspection records indicated that, in all probability, the specifications on the upset end were too tight. A statistical analysis of the hot-punched holes and of the upset squares showed that, if the square were upset to an average 0.005 inch smaller than the average size of the hole, the assembly could be made with complete satisfaction, while a tolerance range was allowed on the upsetting operation of 0.008 inch. The engineering department, upon being shown these findings, agreed to an experimental relaxation of the upsetting tolerance range from 0.004 to 0.008 inch. After a week's trial had shown no defective assemblies, the engineering department agreed to make the new upsetting tolerances permanent.

4. Changing the user's ideas. An excellent illustration of this type of improvement is given by Sir Charles Darwin, director of the British National Physical Laboratory. In a paper that he presented to a joint meeting of the Institution of Physical, Mechanical, and Electrical Engineers, which was published in *Nature* May 23, 1942, he gives an example taken from the manufacturer of time fuses for antiaircraft guns. He says,

> Suppose that the lethal area of a bursting shell is such that, if it explodes within a tenth of a second of the set time, it will make a kill. The gunner therefore demands of the manufacturer that he make a fuse with an accuracy of one tenth of a second. The manufacturer works out his method, but finds that, whereas it is easy to get a fuse accurate to one fifth of a second, he will have a lot of trouble to get one with an accuracy of one tenth; and indeed, he estimates that for the same effort and cost he could not hope to get more than a quarter as many fuses with an accuracy of one tenth of a second as he could fuses with an accuracy of a fifth of a second. Then, half his shells will burst within the range asked for, and so, in fact, he would be wise to accept an inferior fuse, since he would thereby get four times as many shells, of which half would do what he wanted, and he would therefore double the rate of killing. I need not say that I have over-simplified the business; on one side I have left out the cost of the other parts of the shells, and on the other, I have forgotten that the gunner has uncertainties of range to consider, so that his demand for a tenth of a second is more exact than he can justify for practical use.
>
> My example is intended to show that it is good business for the user and the maker of any article to get together before deciding the tolerances of manufactures. The user may be inclined at first to feel that in doing this he is surrendering some of his freedom of choice, but if you will consider it closely you will see that this is not so. He does not have any real freedom of choice, since he must surely try to design the article so as to be as easy to make as possible. I may summarize this aspect of the matter by saying that the user has in the past tended to demand that everything should be made for him as well as possible; but he ought to want everything made for him as badly as possible, or perhaps not quite that, but as badly as permissible. It is in this aspect that statistical quality control especially gives the right information.

The user of a product certainly should not pay for any better quality than he needs. Closer tolerances, finer finish, better fit, more durability —all these, if they exceed what is economical from the purchaser's standpoint, make the product he buys more costly to him than it should be. Sometimes, as Darwin points out, the buyer of a product actually does not know what he should have. Sometimes the manufacturer, by means of education and discussion, can show the purchaser that the thing that is most economical for him (the manufac-

turer) is also economical for the user. If this cannot be done, it is certainly true that the manufacturer, by getting the desires of the purchaser or user in concrete definite form, gains highly valuable information for use in setting his own standards. In any case, the manufacturer has nothing to lose and much to gain by the type of consumer research that will enable him to produce his product more intelligently and economically.

SOME AXIOMS OF QUALITY CONTROL[1]

In order to achieve maximum efficiency through the methods of statistical quality control it is necessary to fit the organization and procedure to the situation in each particular plant. No matter how various the products may be within any plant or between different plants the basic technique of control charts runs like a common thread through all the applications of statistical techniques to manufacturing processes. Successful use of control charts requires adherence to a few fundamental principles that are so basic that they almost can be called axioms.

Axiom One. *Wherever like products are turned out in quantity, control-chart techniques are applicable.*

The question recently was asked of an authority in this field, "What is the smallest number of similar articles on which control-chart techniques can be used?" His answer was, "Two or more." The author knows of one plant in which very large and complicated machines are built, possibly not more than a dozen a year, where control charts have been used very successfully. Usually, however, mass production calls for the manufacture of a large number of consecutive, presumably similar, articles. The remark often is heard, "I can see the value of quality control to some other business, but my business is different." True, every operation is different from every other operation. Nevertheless, the statistical principles that underlie control charts are so universally valid that rarely is found an operation that is not susceptible to this type of analysis. In any plant where scrap, reworks, personnel problems, consumer complaints, engineering design, or costs cause trouble, the control-chart technique is usually applicable.

[1] The author is indebted to William A. Kerr, formerly sales manager of the Spencer Lens Company, for many of the ideas that follow. They first appeared in *Industrial Quality Control*, July 1944.

Axiom Two. *Variability exists in every repetitive operation.*

No two things are ever alike. They may seem similar, and even by every test they may appear to be identical, but they are nevertheless different in some respect, even though our measuring instruments may be too inexact to discover the difference. Such exactness does not appear ordinarily in manufacturing operations. Usually the differences between similar articles are large enough to be measurable, and to require upper and lower tolerance limits that may be either set by the design or followed as factory practice. Suppose that in a machining operation two consecutive pieces vary by 0.001 inch. What was the cause of that variation? It may be impossible to separate out *the* cause; in fact, there may be a multitude of causes that contributed to the difference. Therefore, it may be impossible to assign any one cause or any two causes or any finite number of causes to the 0.001-inch variation. If the usual variation of the process is 0.003 inch, it probably will be true that no cause or causes can be said to have produced the 0.001-inch difference. In such a case we say that there is no *assignable cause* for the variation. Suppose again that the difference between two parts in a machining operation was 0.010 inch. If the normal variation of the process were 0.003 inch, a definite and single cause for that excessive variation probably could be found. It might be, for instance, that a drill was changed; or that a new operator came on the job; or that any one of a thousand other things could have happened. This we would call an assignable cause, or a findable cause. In the first instance, where the difference between two successive pieces was only 0.001 inch, the only thing we could say would be that the "constant system of causes"—that is, all the minor changes in the process that are normal to the process—may have contributed to the 0.001-inch variation. In the other case, however, the variation was so great that in all probability some single cause could be found to account for it.

The control-chart technique is primarily and fundamentally one that enables management to determine what the expected or chance variability in the process is; that is, what the effect of the constant system of chance causes of that operation are. The control chart isolates the excessive variations that it probably would be economical to hunt for and eliminate.

Axiom Three. *Quality must be built into a product; it cannot be introduced through inspection.*

W. Edwards Deming has stated this elementary principle very aptly: "Not how *much* product, but how much *acceptable* product is what counts."

So obvious is this statement that it may seem needless even to mention it. To executives, managers, and planners, who have time to think, and whose job, in fact, is planning, it will be unnecessary to amplify that remark. The working force in a manufacturing plant, nevertheless, facing the hurly-burly of production, meeting and struggling with the daily problems, being pressed with demands on all sides for their time and their attention, often are unable to stop and plan things through. Most foremen and even some department heads, in fact, do not consider it their duty to stop and ponder too much. Their duty is to get things done when things need to be done. A correct division of manufacturing responsibility should place the thinking job on the shoulders of executives and planners, while the working job should be in the hands of those who have the time and the qualifications for their work. What the control chart contributes is a method by which the plant problems that require thinking can be brought to the attention of executives; and, equally important, a method of keeping the jobs that do not need thinking away from them. It also enables the thinking to be done clearly and quickly so that instructions can be issued promptly to the working force. Essentially, the control chart, by bounding the area of chance variations in a process, and by calling attention to variations that are larger than normally should be expected by chance, brings to the attention of management at each level only those problems that need to be considered at that level.

In this way quality—that is, characteristics that meet management standards—can be built into the product during the process of manufacture. Otherwise, defects may pass unknown and unrecognized through the manufacturing process to the final inspection test, and there the product may be rejected, or returned for extra and costly reworking, or even scrapped. Furthermore, a process that can be depended upon to produce good work gives a far stronger guarantee of acceptable quality in the finished product than does any amount of final inspection.

Axiom Four. *A state of control is not usually found.*

When an operation produces articles that remain consistently within their range of chance variation, so that no assignable or findable cause is present, the operation is said to be in a state of statistical control. It is a fact that no longer surprises quality-control men but does sometimes surprise others, that, when a control chart is first applied to an operation, the operation is rarely found to be in a state of statistical control. Even if the inspection methods are superior and the records are complete and adequate this failure or rather inability to analyze

the records by statistical methods, usually results in assignable causes being left undiscovered. Many years of experience, in many plants, by many control-chart men, confirm this fact: assignable causes of variation are present in most manufacturing operations, but in the absence of control-chart techniques they may not be recognized. This circumstance is the basis of the universal need for statistical quality control.

Axiom Five. *A state of control must be established at a satisfactory level before maximum efficiency in the operation can be obtained.*

When, by use of the control-chart technique, the natural variability of any operation has been reduced to the point where a state of statistical control exists, that operation has achieved its maximum efficiency under the conditions as they are at that time. If the controlled process does not produce a satisfactory product, however, the process still remains uneconomical. If, for instance, an assembly job requires a certain type of manual dexterity, and if one of the assemblers *consistently* produces only half of his assemblies in a satisfactory condition, his work is unsatisfactory although his control chart may show a state of *control* with an average of 50 per cent rejections. Such a condition can be interpreted to mean, considering his skill, training, working conditions, home life, temperament, and all the other factors that go to make up his working habits, that he cannot be expected to do any better work than an average of 50 per cent defective. Some basic change must be made in the conditions under which he works. Perhaps he needs more training, in which case special attention should be given to increasing his skill. Perhaps he has private financial problems that distract his attention from his job. Perhaps he is working under bad light. Perhaps he has some physical defect that prevents him from acquiring the necessary manual dexterity. Perhaps he is temperamentally unfitted for the job. Every effort should be bent toward a basic improvement in his working habits; or, if no improvement can be made, someone else should be found to take his place.

What has been said of a man is also true of a machine. Some machines seem to have personalities of their own. If a machine seems not to like the man who is working on it or the raw material that is going into it, or if it has a bad tool or a loose cog, one of the ways in which it can reveal its troubles is in the quality of the work that it produces. If the variability of the machine's work is greater than it should be, as determined by control-chart analysis, there is something wrong, something that probably can be found and corrected. Then and only then can attempts be made to bring the operation of the

machine into line with management-set standards. If, for instance, the pieces produced are consistently too large, there may be a flaw in the set-up instructions, or in the engineering, or in the raw material. Such a defect probably will not be a minor one that can be corrected simply by an adjustment of the machine. Bringing the controlled product of a machine into line with management standards may require hunting for a new raw material; or training a new set-up man; or changing the design. In any case it is very likely that the change will involve a major alteration of the circumstances surrounding the machine's work. This does not mean that the change will necessarily be hard to make. It may be a very simple change, but nevertheless one that will comprise a basic alteration in the machine's working conditions.

When any operation has attained a state of statistical control, valid predictions can be made as to the future product that will come from it; and, when the basic conditions surrounding the operation are such as to produce a product satisfactory to management, then the goal and objective of quality control has been reached. In this book no attempt has been made to cover the whole field of statistical quality control. The emphasis has been upon the basic technique—the heart and soul of the method—which is the control chart. The author has attempted to demonstrate how the control chart works, why it works, and what results can be expected from it. The primary purpose is to show how this simple and effective method can be used to augment manufacturing profits.

CHAPTER 1

THE ROLE OF INSPECTION

Scientific inspection is still an industrial youngster. Only 150 years ago Eli Whitney invented interchangeable parts in assembling guns for the United States Army. Gauges—first the "go" and then the "no-go"—did not come into general use until after the Civil War. Since then, as manufactured articles have become more and more complex, the part inspection has had to play in the production of goods has grown steadily in significance. Nowadays the inspection department, supplied with special equipment and trained personnel, is recognized as being on a par in authority with the manufacturing departments. Inspection is the watchdog of quality.

In spite of the growing realization that inspection is an integral part of production, an equal on the business team, the inspection operation rarely is used to its fullest advantage. The reason for this is partly that, prior to World War I, there was little if any theoretical foundation laid upon which the best use of inspection data could be built.

The Bell Telephone Laboratories were probably the first to start laying this foundation. A beginning was made when W. A. Shewhart conceived the idea of applying statistical sampling and frequency-distribution theory to manufacturing processes, and in co-operation with H. F. Dodge and others elaborated the mathematical theory as well as the practical uses of statistics in manufacturing and inspection. During the quarter of a century between the two world wars this idea slowly ripened to maturity. World War II gave it a tremendous impetus. In recent years statistical theory and practical applications have marched hand in hand with a broader concept of the purpose of inspection, until now no business can afford to overlook the implications of the new technique.

In this chapter is presented briefly the modern idea of what inspection is and what it can do.

PURPOSES OF INSPECTION

When, in the nineteenth century, business men first began to realize the fact that a machine never produces two pieces exactly alike in

shape, finish, or dimensions, and that not every piece produced is necessarily a good piece, their first thought was to set up a method of segregating the good from the bad. From this idea, inspection was born. Many small businesses and some large ones in the United States today have never grown beyond the idea that the function of inspection is primarily to segregate usable from nonusable articles. As a result, some plants keep no records of inspection operations; in many others the records, if any, are fragmentary or not in proper form for analysis; and, even where good records are kept, they sometimes are not used so fully as they might be. Frequently it happens that no comprehensive study has been made of what items should be inspected, for what defects, and by what method. If these shortcomings are true of so-called final inspection, they are even more common in process inspection, where, because the complete segregation of good from bad is usually not feasible, effective inspection may not exist at all. Process inspection, not for the purpose of removing bad work, but primarily for the purpose of finding and *preventing* bad work, is an essential of profitable manufacturing today. Final inspection then can be merely a double check on the successfulness of the process inspection.

No two people will quite agree about the purposes of inspection. The American Standards Association has suggested the first two purposes stated below, to which may be added a third. They are given in order of importance, but will be discussed in reverse order.

1. To decide what to do about *this process or method of manufacture.*
2. To decide what to do about *this batch or lot of product.*
3. To sort the good from the bad, ostensibly so that a *perfect* outgoing quality can be maintained.

TO SORT THE GOOD FROM THE BAD

Sorting inspection [1] sometimes appears to be necessary. Many wartime products are really critical: the propeller of an airplane, for instance, or the bolts that fasten down a gun mount. The Army Ordnance Department realized this when they set up their classification of defects. "Critical," by their definition, implies the existence of a defect in the product that will endanger life or valuable property if it causes failure during use. Under the Army Ordnance plan all critical material

[1] Sorting inspection is a term used here to describe what is commonly meant by routine 100 per cent inspection. It implies that every piece produced is examined and classified as either acceptable or not acceptable with regard to a given quality characteristic. If sorting is performed in a plant it may be at any stage of the process, but is done most frequently as the last act of inspection, just prior to storage or shipment of the product.

must be inspected by being sorted. Critical defects, however, are few, even in times of war, and even fewer when the product is exclusively for civilian use.

Sometimes sorting is impossible, as in destructive tests of armor plate. When inspection is destructive it is especially necessary to apply scientific principles of sampling inspection so that quality can be assured and volume of production will not be reduced appreciably. Colonel Simon,[2] at the Aberdeen Proving Ground, has done brilliant and inestimably valuable work on this problem, but others have not always fully realized its implications. If scientific sampling inspection maintains required quality on critical items in destructive tests, *why will it not achieve the same result in nondestructive tests?* Why should a successful technique be limited only to cases where it *must* be used? Why not extend it to all similar cases where it *can* be used?

These questions are all the more pertinent because of the proved failures of sorting inspection. If sorting is done rarely and only for locating truly critical defects, and if it is performed with extreme care by skilled inspectors to rigid objective standards, with adequate checks on the results, then it may permit only one defect in a very large number of pieces to get by. But, if sorting inspection is used for *all possible defects in the entire factory output,* it cannot be relied upon to accomplish consistently what it is supposed to accomplish and most probably will fail in its objective of maintaining perfect quality. The reasons for the failure have been investigated frequently and have been found to fall into two classes: economic and psychological.

To get good sorting inspection for the entire output of any plant, no matter how simple the product may be, requires, first, skilled and experienced men; and second, an adequate number of them. The cost of these requirements is usually prohibitive; therefore management is prone to cut down the cost of routine 100 per cent inspection and thereby to cut down its efficiency. A case from the author's experience will illustrate the point:

The A. B. Company was a war baby. The owner and manager, who had had a small machine shop before the war, received a large Army order for a very critical part, for which he had bid on a fixed-cost basis. Shortly after he started to make the part, he found it necessary to inspect all his work 100 per cent, and to increase his inspection force to a ratio of about one inspector for every three workers, in order to maintain Army specifications. For a while thereafter, he had very few rejects, and won the Army-Navy "E," but he soon discovered that the

[2] Simon, Leslie E., *An Engineers' Manual of Statistical Methods*, John Wiley & Sons, New York, 1939.

inspection was costing him exorbitantly. He watched the inspectors for a couple of weeks and found that some of them did not seem to be working very hard. He therefore laid off half of them, expecting the other half to do all the work. Only a few weeks later the number of his rejects began to soar, and in spite of strenuous efforts he could not get it down again. He finally lost the contract.

Here enters the psychological aspect of the problem. Sorting inspection is repetitive, in that it usually requires doing the same thing over and over all day long thousands of times. If inspection were simply a matter of pushing a foot pedal or pulling a lever, there would be little trouble; but there is hardly one inspection operation, whether with micrometer, with dial or plug gauge, or with special fixtures of any kind, where human judgment is not necessary. Generally the mental caliber of inspectors should be higher than that of manual workers or simple machine operators. The effect of repetitive fatigue upon the mentality increases in deadliness as intelligence increases; hence repetitive fatigue deadens the judgment and reduces alertness. The inspector gets bored. He is then in no condition to make correct decisions about the acceptability of any particular piece. He tends to take the easiest path and let it pass.

Repetitive inspection leads to lack of mental concentration. The inspector who performs an operation a few thousand times becomes adept at it and therefore careless. He is apt to seek distraction in any movement or talk around him, while his hands and his eyes do automatically what they have been taught. Therein lurks danger, for eyes and hands may slip if not guided by conscious attention. Anyone who has tried speed typing will agree that, if the attention wanders, speed may not decrease very much, but errors will increase. The same thing happens when an inspector who is too sure of his skill allows his attention to wander from his work. In other ways, too, carelessness rears its head. When the inspector gets into the rhythm and continues long at the task, he is apt to neglect the checking of his instruments until by chance something breaks the rhythm, or he gets tired and takes a rest. In the meantime, if his gauge is bad or his measuring instrument needs adjustment, many unacceptable pieces may have been passed through. Furthermore, if the inspector has to keep at the same routine all day long, his standards may change. In the morning, fresh from a night's sleep, he may be too strict and careful in gauging or in visual inspection; as the day wears on he unconsciously may become a little easier in his requirements. Or, if he starts the day's work with lax standards, he may tighten up later on. In either case the lapse of time may make him forget the way he started.

These statements apply particularly to the many inspection operations in which the inspector has to exercise judgment. An experiment that can be made in almost any plant will confirm the dangers of routine 100 per cent inspection. In one such experiment, 1,000 pieces were sent to sorting inspection for polish defects; 600 were accepted and 400 were rejected. A couple of days later the *accepted* pieces were sent back to be sorted, without the inspector knowing that they had previously passed through his hands; this time he rejected 153 out of the 600 pieces. Next, the 400 rejects were reinspected, and the inspector accepted 98 of them. Each time the inspector was offered a lot that he had previously inspected, he changed his mind about what was acceptable and what was not.

Such an experiment can be continued until the positively bad pieces and the positively good pieces have been separated; but frequently more than half of the original lot will have been both accepted and rejected at various times.

Another experiment will illustrate the ineffectiveness of routine sorting inspection. In a certain operation four inspectors examined every piece produced for visual and dimensional defects and threw out the defective pieces. The sorting inspection seemed to be necessary, because this particular operation affected many succeeding operations. It was proposed that sampling inspection replace the sorting inspection; and the quality-control group was asked to investigate and determine whether or not the sorting was effective.

From one lot a random sample of 500 pieces was taken and inspected carefully (as all samples should be): 26 defective pieces were found. This result was analyzed by using Simon's I_Q charts.[3] According to those charts, the existence of 26 defectives in a sample of 500 indicates that (nine chances in 10) the quality of the whole lot should be somewhere between 7.8 and 4.1 per cent, with a most probable quality of 5.4 per cent. Converting these quality limits into number of defective pieces gave

7.8% × 20,000 = not more than 1,760 defective pieces in the lot
5.4% × 20,000 = most probably 1,080 defective pieces in the lot
4.1% × 20,000 = at least 820 defective pieces in the lot

Thus, the sample indicated that there were at least 820 defective pieces in the lot. The four sorting inspectors were not told about this conclusion, but were instructed to sort out the defective pieces in the usual manner. They threw out only 720 pieces. Several other orders were treated in the same way, with the results shown in Table 1.

[3] Simon, Leslie E., *ibid.*

TABLE 1

(1) Order Number	(2) Number of Pieces Produced	(3) Minimum Number of Rejects Calculated from Sample Inspection	(4) Number of Rejects Removed by 100 Per Cent Inspection	(5) Per Cent Efficiency (4) ÷ (3)
W21	20,659	431	248	57.5
T159	28,470	54	39	72.2
W103	15,620	107	61	57.0
F8	5,545	4	4	100.0
K65	10,338	90	51	56.7
B240	10,200	202	95	47.0
V16	11,000	123	59	48.0
Totals	101,852	1,011	557	55.2 average

Although the sorting inspectors were given every advantage, it was discovered, as shown in Table 1, that they found on an average only 55.2 per cent of the *minimum* number of pieces they should have rejected. If the *most likely* number of bad pieces had been taken as the criterion, they would have found only 37.3 per cent of what they should have. Often sorting inspection will give better results than those cited here, but in almost every case it will be found that sorting inspection falls more or less short of its purpose. That purpose usually is *to remove every bad piece.* On the other hand, scientifically designed and correctly performed sampling inspection can give any desired assurance of quality short of perfection. This is true, because, in sampling inspection by quality-control methods, the inspection records are so designed that they can be analyzed, compared, and cross-checked with each other. When this is done, it becomes difficult for any inspector to get away with careless work. When he knows that he is being checked, the inspector is apt to do the best he knows how. Furthermore, unskilled inspectors can be singled out and trained; trouble spots in the process, that result in bad product, can be discovered; the method of sampling, so that a random sample is obtained, can be controlled; and many other aids to efficient inspection are possible.

In the case previously mentioned, where sorting inspection had failed to do the job with four men, scientific sampling inspection did it as well as it needed to be done, with two men—a 50 per cent saving in man power and in dollars.

Nothing in this world is perfect. If sorting inspection is used on the entire output of a plant or an operation with the aim of finding and discarding *every* defective piece, it cannot succeed in the long run. It is expensive; the pay may not be attractive to able men; and the work load may be heavy. Fatigue and boredom may reduce efficiency. For these reasons routine sorting inspection gives no *known assurance* of outgoing quality.

A detail inspection, treated as a special case of a 100 per cent sample, used on lots that fail to pass a sample test, and subject to checks of efficiency, usually will remove practically all defective pieces from the material inspected. It thus will fulfill its objective for most practical purposes. Such inspection, however, is generally too costly and too slow for use on the entire product.

In order to maintain with known assurance a desirable, economical, and profitable quality of product, it is better to admit that nothing is or can be perfect, that the product when it comes to inspection is defective, that it is impossible to remove *all* the bad pieces, and that therefore the outgoing product necessarily will contain some proportion, however small, of unacceptable merchandise. From this standpoint the purpose of inspection becomes quite different. It becomes the problem of how to maintain the outgoing quality (whether outgoing from the plant or from one department to another) at a *desirable economic level.* The most profitable quality is one that compromises between two conflicting interests: the factory's desire for low-cost output, and the sales department's desire for perfect quality. When management has decided upon the most profitable quality level, statistical sampling methods provide the best available means for maintaining it.

ACCEPTANCE SAMPLING

As has been pointed out, routine one hundred per cent inspection, when used in a wholesale way for maintaining quality, is usually a poor method, unreliable, and costly. But detail inspection (a 100 per cent sample) has a definite role to play in any scientific inspection plan. The Dodge–Romig tables illustrate how detail inspection *as a special case of sampling* is occasionally necessary in order to hold a manufacturer's outgoing quality at a desirable level.

In 1941 H. F. Dodge and H. G. Romig of the Bell Telephone Laboratories published in the *Bell Technical Journal* their now famous "Single Sampling and Double Sampling Inspection Tables." [4] Their plans

[4] Now in book form, "Sampling Inspection Tables," published by John Wiley & Sons, 1944.

will assure a manufacturer that the average outgoing quality of his product, expressed as a per cent defective, is at or below any practical and desirable level, and that he can maintain it with a minimum inspection cost. The manufacturer must know (*a*) how large a lot is to be inspected, (*b*) what the average quality of goods submitted to inspection is, and (*c*) what average outgoing quality *limit* he wishes to hold. Then, from the tables, he can determine what size of sample to take and how many defective pieces he can allow in his sample. If the number of defective pieces found in the sample is not larger than the number allowed by the table, the whole lot is passed. If more than the allowable number of defectives are found in the sample, then the whole lot must be detail-inspected and the bad pieces thrown out. This amounts to taking a 100 per cent sample of lots that fail to pass the test; and the detail inspecting must be done with the same care and accuracy as the sampling is done. The manufacturer then can be 90 per cent sure that his average outgoing quality is better than the desired quality *limit* (on the average it will be twice as good), and that the inspection man-hours required are a minimum for the plan he is using. A guaranteed quality at minimum cost is the result.

Another scheme, somewhat similar, is used by the Army Ordnance Department. Given the lot size and average quality submitted to inspection, these tables show the sample size and the maximum number of defective pieces allowed if the average outgoing quality is to be held at any one of several desired limits; they also provide for reduced inspection when 20 successive samples of any one product from any one manufacturer have passed the sampling test. Samples that fail to pass the test require that the whole lot from which they came must be detail-inspected.

It must be realized that no ready-made inspection scheme will fit every requirement of every situation. The manufacturer who wishes to use a scientifically designed inspection system should adapt to his special needs the plan that serves him best. To do this successfully requires a good deal of statistical knowledge and practical experience. If he does not have that knowledge and experience available in his organization, he should find it outside, perhaps on a consulting basis. Expert advice in setting up and installing a sound inspection plan not only is inexpensive in the long run but also is necessary for consistent, reliable results.

CAUSES OF BAD WORK

In process inspection, sampling is the most practical way of determining the quality of the process, because when people or machines are

turning out parts by the hundreds or thousands an hour, routine 100 per cent inspection is usually too expensive. In sampling at the machine, it is very easy to make the serious mistake of not inspecting enough. Although many types of process defects can be determined by the examination of a single piece—such as some errors in setup or certain kinds of maladjustment of the machine—many other kinds of bad work show up only sporadically. Typical of the setup type of defect is the wrong size of drill on a drill press; mismatch in a forging is an example of maladjustment in a machine; occasional or sporadic faults often show up in handwork such as polishing or grinding. In the case of an automatic screw machine, for instance, it is quite possible for pieces taken in succession, one off each spindle, to vary by as much as 0.002 inch in a one-half-inch diameter. If the tolerance were 0.498 to 0.502 inch, and the inspector found one piece at 0.5015 inch, it might be that the machine was producing some parts as large as $0.5015 + 0.0020 = 0.5035$ inch. Again, if the inspector measured two pieces and found one of them 0.502 inch and the other 0.5015 inch, he might be inclined to say that the job was running satisfactorily, whereas actually the work could be seriously out of tolerance, because one of the other spindles might be producing parts as much as 0.002 inch larger. The thing to remember, a simple thought but one that is prone to be overlooked, is this: the pieces that *are* inspected hold significance only because of what they may tell about the pieces that are *not* inspected. Just because all the pieces that are *measured* by the inspector are within tolerance, it does not follow necessarily that all the pieces *produced* by the machine are also within tolerance. For this reason, sampling process inspection must be done in such a way that sound conclusions can be drawn about the quality of the entire output from evidence provided by the sample. If the sampling is done in the wrong way, or inadequately, control of the process may seem to exist while actually it does not.

Another kind of bad work can be traced directly to workmen, especially when the operation involves manual labor: grinding, for instance. Because of poor training, lack of dexterity, hurry, anxiety, fatigue, or carelessness, bad grinding can occur often, in fits and starts, during a day's work by an operator. Under such circumstances inspection of one piece, or a few pieces, or even many pieces off the top of the pile of work usually will fail to reveal the full extent of the trouble. Only an adequate sample, taken from all the work done since the inspector's last visit will tell the true story. How large a sample to take, how frequently, how to get it at random, and how to

interpret what is found in terms of process quality, constitute one of the contributions of statistical quality-control technique to process inspection.

WHAT TO DO WITH THE LOT

When a company installs statistical sampling inspection, the management may discover that the quality of their product is not so good as they had thought. They also may find that their manufacturing process is not in control. As a result the outgoing quality may be spotty. Sometimes the product is good, and at other times poor. The poor spots are usually unpredictable, both as to time and as to cause. Thus inspection is posed its second problem: What to do with any particular batch, lot, or run of the product.

In this respect the C. Company was typical. The company manufactured a variety of parts and subassemblies for airplanes, most of them in large quantities. Before starting a quality control program, the management decided to investigate the need for it. By sampling a half-dozen different products after they had gone through final inspection, the company discovered that there were almost 3 per cent defective parts going out of the plant. None of the defects, fortunately, was critical, but the outgoing quality was nevertheless undesirably poor. The faults fell into three categories: bad raw material, bad machining, and bad finish. The worst batch was one where 30 per cent of the pieces had the wrong part number marked on them. A few cases were found where two parts were assembled wrong-end-to, and even more frequent were parts on which one or more operations had been omitted.

Investigation showed that most of this bad work was caused by carelessness in the plant. Very significantly the reason such defects got by the final one hundred per cent inspection was that they were so obvious and unexpected that no one looked for them, or looking, no one saw them. This was not so impossible as it sounds, because the inspection was so organized that certain inspectors looked for certain types of defects, but those that no one looked for (because they were not anticipated) sometimes got by.

As a first corrective measure the company instituted a "pilot" sampling inspection just before each lot of goods was finally inspected. The sample told them what most frequent defects to look for in the sorting inspection and pointed to the trouble spots in the plant. A systematic check and follow-up, proceeding from the pilot inspection in one direction to the outgoing product and in the other direction back into the plant, was begun. Statistical techniques were used (*a*) to check the adequacy of the sorting inspection by comparison with the

"pilot" findings and occasionally by taking a sample after as well as before final inspection; (*b*) to put control charts on the erratic operations in order to prevent future recurrence of the same fault. Depending upon the "pilot" inspection to tell them in advance what to look for, the C. Company is gradually making the transition from sorting to sampling inspection for all outgoing product. As quality becomes stabilized, inspection becomes less expensive and more efficient. Final inspection man-hours are now only half of what they were before the program was begun.

The question, "What to do with this lot?" now answers itself almost automatically. In the C. Company's present setup, one of three things can be done with a lot at final inspection.

1. If the lot passes the "pilot" inspection, it is shipped out immediately without being sorted.

2. If the quality of the lot falls within a range where it is economical to sort it and to scrap or rework the defective parts, the lot is submitted to detail inspection for removal of the bad pieces, and a sample inspection is made after the detailing to be sure that it has been done efficiently.

3. If the quality of the lot is so poor (usually because of an "unexpected" failure in the plant) that sorting it is too costly, it is rejected and returned to the factory for repairing or scrapping.

The criteria used in determining which of the three dispositions shall be made of the lot are economic ones. If the lot fails to pass the pilot inspection, and the second or third type of action has to be taken, a quality committee composed of representatives from the sales, production, and inspection departments decides what shall be done.

Somewhat different was the problem faced by the D. Company. All its work was done by about ten subcontractors. Even assemblies were made on the outside. The D. Company, therefore, was concerned primarily with production scheduling and inspecting. It had two contracts for critical and urgent war material, but found that, because the subcontractors could not be depended upon to produce the required quality of work, schedules were lagging. Often badly needed parts had to be rejected because of critical defects, and thus assembly of ten times as many good parts was delayed.

Fortunately, the D. Company had complete inspection records for six months back. Statistical analysis of these records told them which subcontractors most frequently made which mistakes.[5] In some cases the subcontractor did not have the right equipment to do the job; that

[5] See Chapter 5, Case History V.

particular work was therefore transferred to another subcontractor who could do it; if possible, the original "sub" was assigned a job for which his machinery was adapted. In one or two cases it was the subcontractor's inspection rather than his process that was at fault; help then was given the subcontractor to improve his inspection. A control chart was kept for each job done by each subcontractor, which showed his percentage of bad work lot by lot as it was received by the D. Company. Frequently receiving inspection was done by field inspectors at the subcontractor's plant, and charts were kept right there. Whenever trouble showed up on the charts, the subcontractor was asked for an explanation. Every month each subcontractor was sent a photostatic copy of his charts with full explanations and if necessary special analyses accompanying them.

This work had two immediate results. First of all, it brought about a complete reshuffling of subcontracts and of scheduling. The D. Company, knowing exactly what kind of work each "sub" could do best, was able to plan its own schedule with a higher degree of assurance that the required quantities would be delivered at the right times. Secondly, the D. Company found that its own inspection was easier and less expensive, because much less rejectable work was being submitted. One dimension of a gun mount that formerly had been given complete inspection with micrometers was accepted on the basis of sampling. Another part that had previously undergone a 10 per cent destructive test now was passed with only a 2 per cent test. Besides these and other savings in time and money, the D. Company found its relations with its subcontractors improving so rapidly that within a few months it released more than half of its expediters and cut its clerical force by one fourth. The inspection force, however, was increased. The extra man-hours were used to develop more exhaustive tests, more accurate gauges, and improved inspection fixtures; where sampling was done, larger samples were taken and the management's degree of quality assurance was correspondingly increased.

Both C. and D. Companies realized, even before they had started their statistical quality-control work, that finding the defective product at the last stage, at final inspection, would do some good but would not solve any problems permanently. The findings at inspection had to be translated into terms of production processes and operations. Thus both companies were led directly to put into effect the most important of all the purposes of inspection. That purpose is to stop bad work *before it is produced,* to prevent defects rather than to hunt for them. Inspection fulfills its highest function when it helps to make things right in the first place.

WHAT TO DO WITH THE PROCESS

If machines and people could make two consecutive articles exactly alike, there would be no need for statistical quality control, nor for inspection. But in manufacturing, as in all things, the one certainty is that any two articles produced under presumably identical conditions will be different. Interchangeable parts, the basis of mass production, are made possible only by keeping the variations so small that they do not matter in practical use. It is clear, therefore, that a successful process is one that can maintain *consistently* the required limits of variation.

A process is a series of operations designed to produce an article with certain desired characteristics. Each operation in the process must fulfill its own specifications in order that the process may be successful. These specifications, it should be understood, are not necessarily the requirements set up by the engineering department, but are those that are necessary to make the finally completed article acceptable as to cost and salability. Sometimes the operation is able to meet the demands made upon it, sometimes not. If not, management must know, so that a different operation can be designed that will meet the requirements; if, on the other hand, the operation is capable of doing the job it is supposed to do, management must know whether or not it is doing that job. In either case the variations in the operation must be controlled at a level that represents an economic compromise between the cost of production and the quality of product.

This is the great contribution of statistics to manufacturing: that only by statistical techniques can variations be analyzed qualitatively. Since the essential requirement of mass production is the control of variations in the product to such a size that they do not interfere with each other when the product is put to use, any technique that will help toward this end is desirable; and, if that technique is basic to the *control* of variations, it is not only desirable but also necessary.

Each operation in a process is a small world of its own, where the laws of cause and effect rule just as they do in the whole universe. Variations in the product—in size, shape, position, hardness, strength, finish, and so on—are the effects; the causes are raw materials, tools, dies, setups, machines, workers, supervisors. These causes act and interact on and with each other in a tangled maze, producing articles whose characteristics vary from one piece to the next. The manufacturer wants to know *in advance* whether these characteristics will stay within the limits assigned to them; that is, he wants to be able to predict correctly the limits of variation in the process.

Every day, and every minute of every day, people act on the basis of predictions. If the predictions are correct most of the time, the predictor is said to have good judgment; if they are bad, he has poor judgment. A man probably would be willing to risk his reputation on forecasting that the sun will rise tomorrow morning. But to predict that he will *see* the sun rise tomorrow morning is another matter, because he then will have to make at least two other implied forecasts: that he will be there to see it, and that there will be no clouds. He cannot be as confident of the second prediction as of the first. If he tosses a single coin, he cannot predict whether heads or tails will come up. But, on the other hand, he should be more willing to state that out of 100 tosses at least 45 heads will appear; he should be fairly certain to get at least 40 heads; and he can be almost positive that at least 35 heads will appear in 100 throws of the coin.

Applied to factory operations, that is, to the characteristics of the articles produced, sound predictions have a high monetary value. Suppose that a manufacturer makes socket wrenches. If he can predict that the lathe operation will produce an opening that will give a snug fit on every bolt and nut for which the wrench is designed, and if the prediction is correct, it will be worth a lot of money to him. Suppose he can predict correctly that, of the next 50,000 turnings off an automatic screw machine, 1,500 will be outside the tolerance limits: with that forecast in hand the manufacturer can do one of three things with the operation before it starts. First, if 1,500 bad pieces are too many, he can try to improve the process so that probably 100, or 10, or none, will be outside. Second, if 1,500 are too few, he can perhaps let the tools run longer, allowing 2,000 or 3,000 pieces to be outside tolerances, and cutting his costs. Third, if 1,500 defectives in 50,000 is about the right proportion, he can let the operation go as it is. In any case, he can make a decision in advance and can take action in the desired direction.

If, however, the manufacturer is unable to predict correctly what the results of the operation will be, he can take no action in advance. He can only hope that the job comes out all right, that the cost is what he can afford and that the quality is what he can sell. When the run of 50,000 turnings is completed he may find that the cost was too high, or that the quality was too low, or that it was just about right. Such guess work is poor business. It is better business to be able to tell ahead of time what will happen and to have the operation turn out as one wants it to.

When the results of the operation, that is, the quality characteristics of the product, vary in such a way that valid predictions can be made,

the process is said to be in a state of statistical control. With a process in control, economic decisions can be made with a high degree of confidence that they will be correct. If, during the operation, one of the causes gets out of balance with the others (a loose collet or a dull tool, for instance) and produces a greater-than-expected variation in some quality characteristic, a properly designed control chart usually will reveal that fact. Thus the offending cause can be isolated and corrected before much damage has been done. Such an out-of-control cause, as revealed in the product, is known as an *assignable cause.* It must be looked for and eliminated before control is again restored. The *methods* of statistical control, used with intelligence, honesty, and skill to attain a *state* of statistical control, offer strong support to sound judgment and common sense.

CONTROL OF THE PROCESS

Those who have applied statistical quality control agree with those who have built up the theoretical framework, that sampling applied to process inspection is basic to the effective control of quality. What statistical quality control does is to place a new emphasis upon the gathering and analysis of factual data, and thus to provide a new preventive and corrective technique whereby the process can be stabilized. Other control techniques are available to factory management, among them production control, laboratory tests, experimental design, production engineering, and personnel records. The unique function of statistical control is to supplement all the other controls by testing the validity of their data, and to provide continuous graphic clues to sources of trouble in factory operations. Using data gleaned from inspection records, statistical quality control applies scientific analysis in helping to cut costs with one hand while building sales, through higher quality, with the other. Statistical quality control has proved itself to be one of the best ways to make inspection and production profitable. The chapters that follow describe this technique and illustrate, with examples taken from actual experience, how it has been applied to a variety of specific factory problems.

CHAPTER 2

STATISTICAL CONTROL

Many a factory executive, department head, and foreman face each day wondering what crises will arise before the end of the shift. They cannot predict where trouble will strike, what jobs will have to be stopped, how much reworking will have to be done, or when the promised production will be delivered. Many an inspector wonders why defects that should have been found at his station show up later or at other operations, and why material that he has passed as satisfactory is rejected by the buyer. Such mishaps create embarrassment and confusion. They are apt to make the factory man who carries any share of management responsibility wonder whether the job is worth the strain of meeting daily emergencies that seem to strike without warning.

Typical of such problems is that of setups. A drilling operation, for instance, may be set up and approved on first-piece inspection; a few minutes later the job may be stopped by the inspector because it is out of tolerance. A first-piece inspection of an automatic screw machine may show a dimension too large; after adjustment it may seem satisfactory, but soon fails to pass the go–gauge test. Or a hand operation, such as grinding, may run along smoothly for a while and then suddenly develop a large percentage of rejects.

Often during the course of an operation some change occurs that alters the character of the product. A change of drills, for example, may change the size of the hole; a slight adjustment of the tools in a screw machine may change the dimensions of the turnings; a change of operators may cause an unexpected amount of poor grinding to be done. Sometimes a failure of some part of the machine, or a poorly designed jig, or a new batch of raw material, or a new inspector or foreman may be the source of trouble. These difficulties sometimes are found before serious loss has resulted; frequently, however, they are discovered only after a considerable lapse of time, or in some other department, or at final inspection.

MAKING VALID PREDICTIONS

What factory supervisors need, at all levels of responsibility, is some way of telling in advance where trouble will develop and from

what cause. That is, they want to know when a cause of excessive variation enters the process, so that correction can be made immediately, and so that the same *findable cause* will not occur again. They would like to have a process of which they can say, "This dimension can be held within certain definite limits."

Such a statement is a prediction, because it implies foreknowledge of the future. If the foreknowledge turns out to be correct, it is a valid prediction. The ability to make valid predictions is the secret of successful control.

Note that the idea of control, or the making of valid predictions, has no necessary relationship to tolerance limits or specifications. A manufacturing operation can be "in control" in the sense that valid predictions can be made as to the quality of future product, without its meeting specifications; or, if the tolerance range is wide enough, findable causes of excessive variations may exist while the product fully meets specifications. In neither of these cases can the operation be considered as satisfactory: in the former, it can safely be predicted that a certain proportion of the product will fail to meet specifications; in the latter, no valid prediction as to the quality of the product can be made, because, although from past experience the product may meet the specifications, no confidence necessarily can be placed in its doing so in the future. A good operation or process should fulfill both requirements: (1) It should be in a state of statistical control; (2) it should be in control within the tolerance limits. As long as these requirements are met, the valid prediction can be made that the operation or process will continue to produce a satisfactory product.

In manufacturing the emphasis has been concentrated so long on meeting tolerances that the necessity for statistical control, that is, for a state of valid predictability, often has been overlooked. In other fields the need for statistical control is even more apparent, especially in all phases of applied science. Usually a scientific experiment is repeated carefully a number of times under what the scientist calls "essentially the same conditions." How can the scientist know, however, whether the conditions actually are "essentially the same"? No matter how many precautions he may take, he cannot be sure that every findable cause of variation has been eliminated. Furthermore, if his experiment meets the requirement that it shall be made under conditions that allow someone else in another laboratory to duplicate his work, he should make reasonably certain that his results are predictable, before he can be satisfied with his conclusions. The scientist, for instance, may have several assistants working with various pieces of equipment. At the least, he should test each assistant and each

apparatus for any bias that may exist in the observations. Such personal and individual biases are familiar phenomena in experimental work; yet the techniques of *statistical* control rarely are applied to them in order to separate the unreliable observations, if any, from the reliable ones.

In economic forecasting the situation is somewhat different. There the problem is not one of controlling the data, but of deciding when the data indicate the existence of a change in basic economic conditions. The author has applied statistical-control techniques, with some success, to a number of economic indices in order to determine the turning points in business cycles. The author's other applications have been to personnel problems, production control, sales demand, and analysis of balance sheets. Statistical control has been used effectively in the fields of biology, agriculture, fisheries, medicine, and vital statistics. It has become apparent that, in practice as well as in theory, statistical control should not be confined to manufacturing problems. It is a general statistical technique of deep significance and broad scope. It is a basic test for the validity of factual data, useful in a wide variety of problems where decisions are made on the basis of objective evidence.

THE DEGREE OF RATIONAL BELIEF

In every decision, large or small, there are three elements. First, there is the evidence, including both the facts and the interpretation. Second, there is the degree of belief or confidence placed in the evidence. Third, there is the action taken to implement the decision.

Some decisions are almost entirely emotional in nature, as when a man chooses the girl he wants to marry. Other decisions seem rational, but are basically emotional, as when a man votes for his political party's ticket. Still other decisions appear to be based on hunches rather than facts, that is, they draw upon a store of personal experiences for the evidence, which carries with it a high degree of belief. Whether or not such decisions prove to be sound depends upon intangible unmeasurable factors that, in the present state of the social sciences, cannot be evaluated scientifically.

Often, however, decisions can be made on the basis of quantitative facts, scientifically collected and analyzed, that comprise evidence to which a known degree of *rational* (as distinguished from emotional) belief can be attached. Such evidence, coupled with rational belief or confidence, leads to decisions and actions that, although not always right, have a fairly definite probability of being right.

As a contrast of a nonrational with a rational decision, consider a problem that arises frequently in almost every factory. The process inspector at a certain operation insists that out-of-tolerance work is being done and that therefore the job should be stopped; the foreman argues that the operation is no worse than usual and therefore should be allowed to run. The inspector submits as evidence one piece that failed to gauge; the foreman submits five pieces that did gauge. Should the operation be stopped, or should it be allowed to continue?

A decision based solely on the evidence as submitted might be difficult; the referee probably would have to fall back on his general knowledge of factory conditions and his relative confidence in the judgment of the two men. Here the evidence would be inadequate and the degree of rational belief would be low; hence the probability of the referee's making a sound decision might be little better than 50–50, or 0.50.

Suppose, however, that the foreman, instead of submitting five pieces that did gauge, submitted a control chart on the process, indicating that a state of statistical control existed, and that no evidence of a findable or assignable cause was present. Then the job should not be stopped merely for adjustments or other temporary corrections to be made. Any such action almost certainly would fail to improve the operation, because there would be no findable cause for the bad work, or at least no cause that could be found in an economical length of time. Even if the operation, as analyzed by the control chart, were running consistently out of tolerance, it still would be uneconomical to stop it for a minor adjustment of some sort. A thoroughgoing investigation could be undertaken, but a minor shutdown would only lose precious working time.

Suppose further that, as a result of this incident, an extensive investigation was made, and it was determined that 3 per cent out-of-tolerance work was the usual expectation for the operation. With this evidence at hand management could decide, from a study of costs and other factors, whether to permit the operation to run at a 3 per cent level, whether to design some other operation to take its place, or whether to eliminate the faulty operation (and possibly the product) altogether. Here it would be possible to make another sound decision, based upon factual evidence carrying with it a high degree of rational belief. Such a rational decision might have better than a 99 per cent chance of being right, in contrast to the 50 per cent chance of a nonrational decision.

One of the greatest monetary rewards to the manufacturer who has applied statistical-control techniques to his operations—even if they

are not always within tolerances—is that a multitude of minor and vexing decisions that arise every day automatically settle themselves at the lowest level of supervision, without intruding upon the expensive time of higher-salaried executives. Furthermore, the upper tier of management can associate with the data they get through statistical-control methods a high degree of rational belief or confidence, and their major decisions therefore will have a correspondingly greater chance of being sound and profitable.

CONSTANT SYSTEM OF CHANCE CAUSES

Valid predictions and sound decisions can be made only from data that comes from a *constant system of chance causes.* As applied to manufacturing processes, an operation is said to be in a state of statistical control if from the evidence gathered (the sample) it can be deduced with a high degree of rational belief that the operation behaves like a constant system of chance causes.

A constant system of chance causes probably can be defined best by analogy to a lottery drawing from a revolving drum. Assume that all the tickets in the drum are essentially similar, that they are thoroughly mixed, and that the drawing is performed mechanically by a blindfolded drawer. Can any knowable reason be assigned for the drawing of any particular ticket? No, because the selection is a purely chance one. Presumably, an omniscient mind could trace the train of infinitesimal causes and effects that led to the selection of a particular ticket; but, so far as human ability goes, each of the causes is so small, and there are so many of them, that the final result of the drawing is unpredictable. The tickets, the drum, the method of drawing, and all the other appurtenances of the lottery are called a *system of chance causes.* If a series of such drawings is conducted, all under the same conditions, the drawings may be said to be made under a *constant* system of chance causes.

This physical analogy can be applied to a manufacturing operation. Think of the operation as producing, in the past, in the present, and in the future, an indefinitely large sequence of successive articles with a certain characteristic, such as an outside diameter. Think of each article as a lottery ticket. If all the articles are essentially similar as to outside diameter, any samples taken and measured at any time during the operation, when plotted on a control chart, will behave as if they were drawn purely by chance, that is, at random, from a constant system of chance causes. The operation then is said to be in a state of statistical control.

In any operation there are three time elements—past, present, and future. Samples, of course, can be taken only from the present; analysis can be made only from accumulated samples, that is, from past data. Experience has shown, nevertheless, that, when such past data indicate a state of statistical control, the operation has achieved a state of equilibrium or stability that can be expected, with a high degree of confidence, to extend into the future. Hence, when the past data exhibit statistical control, the characteristics of the future product can be predicted in the words, "The characteristic (such as outside diameter) of the articles that will be produced in the immediate future will be essentially similar to that produced in the immediate past." Thus, valid prediction about future quality becomes possible.

An out-of-control operation can be illustrated by returning for a moment to the lottery drawing. Suppose that the even-numbered tickets were larger than the odd-numbered ones, and that the method of drawing favored the large tickets. If these facts were known to an observer, he could predict safely that even numbers would be drawn more frequently than odd numbers. Or, if he did not know the cause, he could deduce from the results of past drawings that (*a*) even-numbered tickets will have a better chance of being drawn in the future than odd-numbered tickets, and (*b*) that there is an assignable, knowable, or findable cause for the discrepancy. If the observer were interested in fair play, he might investigate, discover, and correct the assignable cause. He then would be performing an *operation* of statistical control in order to achieve in the drawings a *state* of statistical control.

During one of the courses in statistical quality control conducted under the auspices of the Office of Production Research and Development of the War Production Board during the war, an amusing instance of lack of control occurred. The instructor was demonstrating random sampling by means of blindfolded drawings from a bowl of red and white glass beads. Ten per cent of the beads were red, and 90 per cent were white. The beads that were drawn each time were replaced, and the beads in the bowl were then stirred around, and supposedly thoroughly mixed, after each drawing. To the instructor's embarrassment, however, far too few red beads were drawn. An investigation showed that, though both kinds of beads were the same size, the red beads contained a mineral coloring agent that made them heavier than the white beads. The red beads therefore tended to sink to the bottom of the bowl, and the white ones, being on top, were drawn more frequently than was expected. In this case statistical analysis gave evidence of a significant difference between the results that actually

were obtained and the results that were expected, leading to an investigation and the discovery of an assignable cause for the discrepancy.

Similar situations frequently arise in the factory. Suppose that process inspection on a certain operation establishes an expected quality characteristic—whether a dimension or a percentage—and then a change of operators, of tooling, of raw material, or of some other factor alters the complexion of the operation, often to an unknown and unrealized extent. The entrance of such changes is what makes many factory processes unpredictable, in the sense that little confidence can be placed in the future product being, within the limits of expected variation, essentially similar to the past product. These unrealized but discoverable changes may or may not cause out-of-tolerance work. Even if they do not result in an unacceptable product, they create an intermittent hazard to costs and schedules, because it never can be known when they might cause bad work and require repairs, machine stoppages, and extra inspection. Statistical control provides management with a tool for determining promptly and inexpensively just when the otherwise unrealized changes occur, how serious they are, and how they can be prevented from reoccurring. Recurring repetitive errors often can be eliminated by the full use of statistical-control methods, which thus contribute largely to increased manufacturing efficiency.

Many other advantages accrue from a statistically controlled process. Personnel efficiency is increased; supervisors are kept on their toes; raw materials are economically used; tooling and machine problems are reduced; estimates and bids can be made with greater assurance; sound predictions can be made; work flow becomes smoother. Above all, the study and analysis of the process develops a wealth of information and leads to many improvements. This alone usually repays many times over the cost of the required engineering and statistical investigations.

RATIONAL SUBGROUPS

Benefits such as these are not gained except by hard work and intelligent co-operation among production men, inspectors, engineers, and statisticians. Since he is a recent addition to the factory staff, the statistician has an opportunity limited only by his ability and breadth of vision. Among the modern statistical techniques with which he must be familiar and which he can use most effectively, is the theory of small samples and its application to process inspection by means of rational subgroups. Classical statistics considered small samples so unreliable as to be valueless. Modern statistics, through the work of

R. A. Fisher, W. A. Shewhart, L. H. C. Tippett, S. S. Wilks, and others, not only has developed a firm foundation in theory for the use of small samples but also has discovered ways and means of turning theory into practice.

Between classical and modern sampling theory there exists the same difference as between a balance sheet and an operating statement, or between a still picture and a motion picture. One is a cross section, the other a flow chart. In the control-chart technique an adequate sample is accumulated by taking small samples of from two to ten observations grouped in a way that has an *engineering basis.* Usually the significant element is time: a small sample of not more than ten items is taken off the machine every half-hour or hour or at some other interval of time. Statistical analysis is carried on *while these small samples are being collected,* instead of waiting until a large sample of 100 or more has been gathered. Sometimes the small samples are taken from lots in order of production, without any regular time interval. It may be that not time but a difference between operators or machines is the significant factor in determining the way in which the small samples are taken. Whatever rational, logical, or engineering basis is used for deciding how the small samples shall be selected, the general principle to be kept in mind is:

Take rational subgroups of small samples in such a way that the suspected assignable causes of trouble can be detected by differences between one sample and another.

As an illustration of this principle, consider three of the ways in which samples can be chosen from a six-spindle automatic screw machine.

1. A large sample of 100 to 500 turnings can be taken from a run after the run has been completed, and frequency distributions can be made and analyzed. This method gives little information on what happened during the run, and why. If 10 per cent of the turnings are oversize, there is usually no way in which the bad work can be traced to any particular machine, operator, or lot of steel.

2. Small samples of two to ten pieces can be taken off each spindle at intervals during the run, and a control chart can be kept on each spindle.

3. A sample of six pieces, one off each spindle, can be taken at intervals during the run, and one control chart can be kept on the machine.

Of the three alternatives the first is poorest because of its inefficiency in providing information for correcting bad work. Both of the other two break the large sample up into rational subgroup small samples;

one of them is therefore preferable. Which one to use depends upon the principal cause of irregularity in the work—an engineering rather than a statistical problem.

An automatic screw machine is basically a machine that cuts and forms metal by forcing a rotating metal bar against one or more stationary cutting tools in succession. Each spindle contains a bar of metal. When a cut has been performed on the first bar or spindle, it moves on to the next tool, while the second bar takes its place at the first tool. Thus each spindle takes its turn at each tool. Differences may exist between the spindles; successive finished pieces come off successive spindles and off different bars, with the result that the product of any one spindle may vary significantly from that of any other spindle for two reasons: differences in the behavior of the spindles themselves (loose collets, for instance) and differences in raw materials. If either of these causes is suspected of producing bad work, sampling method number two should be used because, when samples are taken off *each spindle* and analyzed, the differences between the spindles become apparent. Consequently, if significant differences do exist, they will show up in the analysis.

Several objections to the second method can be mentioned. It is a relatively expensive way of finding out what the process is doing, because a large number of pieces must be measured and six control charts must be maintained. Furthermore, it does not give a comprehensive picture of the behavior of the machine as a whole or what tolerances the machine actually is holding or how a tool is wearing. For such information the third method is better.

If one piece is taken off each spindle—that is, six successive pieces as they are produced—normal variations in the hardness of the metal, the condition of the spindles, and such will be *included* in the sample variation. The third method is valuable if the causes that are correlated with *time* are the suspected ones. These may be tool wear; tool sharpening, resetting, and adjusting; night and day shift, and so forth. A six-piece sample, one off each spindle, taken at intervals of from 15 minutes to an hour and plotted on a control chart will reveal time changes clearly. Exceptional differences between bars or spindles also will be revealed in the range chart, as described in Chapter 3.

Occasionally some method of sample taking other than those already mentioned is required. A battery of automatic screw machines was producing a critical component of a gun-recoil assembly. The usual process-control sampling methods were tried without avail. Every machine individually seemed to be in control at a satisfactory level. As a last resort, a basket was put at each machine in which one hour's

production was placed. A sample of ten pieces from each basket was taken and a control chart drawn up on which inspection results on each machine were shown at the end of the hour. One group of five machines showed an unusually small variation within their samples; either those machines were exceptionally reliable, or the measurements were wrong. Since no reason for their superiority was known, the dial indicator used in measuring the critical dimensions was examined. It was found to be defective, or lacking in sensitivity. The flaw in the measuring instrument had been responsible for over- and undersized pieces, since the inspector accepted all the pieces passing the test. Some of them, however, were actually out of tolerance because the indicator failed to record the extreme measurements.

Making the right decision about the subgroup to be used in process-control inspection calls for the combined knowledge of a good engineer, a practical factory man, and a seasoned inspector. Even then mistakes may be made until, by trial and error, the most sensitive and useful method is arrived at. This part of statistical quality-control work is worthy of long and patient effort, because it pays large dividends when it is successful. The statistician's role at the trial-and-error stage should be to assist in finding the appropriate subgroup method by studying the process and interpreting the results of the various experiments until the right one is found. There is no substitute for control-chart experience in the preliminary stages of such an investigation.

Enough has been said about rational subgroups to point out a few of the factors involved, most of them nonstatistical. From the standpoint of statistical theory, the purpose of subgrouping is to set up data for analysis in groups, the items in each group being considered as essentially similar. Then, if there are discoverable nonchance causes, their effects should appear as excessive variations *between* the samples.

W. Edwards Deming has well stated the principle involved in the selection of the rational sub grouping for a control chart:

> Each point on the control chart is derived from inspection tests carried out on a sample from a batch of product that for engineering reasons is thought to be produced by a common system of causes. The batches are selected in some rational manner, the purpose of selection being to disclose *assignable causes of variability between batches*. A well-planned selection of batches for sampling is extremely important, and presupposes enough knowledge of the manufacturing process for forming intuitive premonitions regarding the sources of variability. The plan for selecting batches for sampling is like laying a trap for catching variability. Often it turns out that small samples selected in order of production are effective.

The points that are plotted on the chart will show variation in the quality of the product, batch by batch, hour by hour. Control limits are placed on the chart. They provide a basis for action, because they discriminate between causes of variability that can be discovered (assignable causes) and causes of variability that cannot be discovered (chance causes). When a point falls outside the control limits, lack of control is indicated, and it will pay to look for the cause. Whether you are manufacturing or buying, lack of control indicates a spotty or variable condition and the need for doing something about it. If you are a purchaser, you will need increased inspection to maintain protection.

SUMMARY

The discussion in this chapter has been necessary to establish the philosophy underlying the control-chart technique and to describe the *state* of statistical control. In the following chapters the *operation* of statistical control is described, that is, the procedure necessary for reaching and maintaining the *state* of control. Without realizing in advance *why* and *how* control charts work, the reader of this book might be tempted, because of the apparent simplicity and ease with which control charts can be used, to draw rash conclusions as to their interpretation. Control charts are powerful analytical tools, but, like all keen instruments, they give best results when used with skill and understanding.

CHAPTER 3

$\bar{X}$ AND R CHARTS

Control charts are of two kinds: charts for variables, in which quality is described quantitatively in terms of dimensions, weights, or other characteristics; and charts for attributes, in which inspection is visual or by go–no-go gauges, with the product classified as either good or bad. In this chapter charts for variables, using the average, $\bar{X}$, and the range, R, are discussed. In the next chapter the p and pn charts for attributes are considered.

STARTING THE CONTROL CHART

Several preliminary steps are necessary before setting up a control chart for either variables or attributes.

1. Decide what quality characteristic of what product on what operation the chart is to be kept. Some feature of the process that has been causing trouble and extra cost in scrap or salvage and that has not yielded to corrective efforts is a good place to start; maximum economic results will be obtained by a control chart in such a spot. Alternatively, if a demonstration of the power of control charts in making improvements is desired, an operation and characteristic that is important, but that apparently has been causing no trouble can be selected. Frequently a control chart on a non-trouble-making operation will reveal the existence of unsuspected trouble and lead to otherwise unattainable improvements. It is usually wise to choose, for the first chart, a place where adequate inspection records already exist or can be collected easily without too much extra work; that is, select a quality on which inspection is inexpensive.

2. Study the process. Ideas should be gathered from operators, foremen, supervisors, inspectors, and engineers. It is advisable to assign the responsibility for developing the chart to a man with statistical training, or to a sympathetic-minded engineer. He should have a thorough practical knowledge of the whole process, of the effects of each operation on the succeeding ones, and of where troubles may exist, before he begins to collect the data. He should acquaint himself with the specifications and with their relationship to shop practice.

3. Study the method of inspecting an individual article or making a single observation on the selected characteristic. Note particularly any factors that may give rise to errors of measurement or of observation. Prepare written instructions to the inspector specifying the characteristic to be inspected; the instrument to be used, such as micrometer, dial indicator, gauge; the method of making the observation; the accuracy required; the size of sample; the time schedule; and the method of recording the data. If possible divide the work into two parts, have a different inspector do each part, and check one against the other.

4. Prepare forms in advance on which the inspection results shall be recorded. Clear instructions should accompany the forms. If the calculations are not to be made by the inspector, specify who shall make them and to whom he shall report. The author never has found an all-purpose form that can be used generally. It is usually better to design special forms for specific purposes: it is less expensive to handle them this way than to do the additional writing-in and clerical work required to make a general form apply specifically to a given operation.

5. Choose the rational subgroup so that, from knowledge of the process, it is unlikely that an assignable cause is operating *within* the sample. For instance, shall a control chart be kept on each machine, each operator, each shift, or each source of raw material; or shall the whole stream of product be treated as a constant system of causes, with one control chart kept on the entire production?

6. Set up a procedure that will enable the inspection results to be charted and analyzed as soon as possible. In control-chart work speed is the essence of effectiveness. A cause of trouble discovered *and acted upon* within 15 minutes is far more effective than one that is delayed an hour. If an assignable cause is present, the more quickly it is eliminated, the easier it is to keep the process in control. Time is money in such cases. Quick action not only corrects production faults promptly but also maintains strong psychological pressure on foremen and workmen to produce good work and prevent out-of-control points before they appear.

TERMINOLOGY

In control-chart work the following terminology is generally used:

m = number of subgroups or samples
n = number of observations in each sample
$N = mn$ = total number of observations.
$\bar{X}$ = arithmetic mean of the n observations in each sample

$\bar{\bar{X}}$ = arithmetic mean of all N observations
R = range in any sample (largest value minus smallest value)
$\bar{R}$ = arithmetic mean of the ranges in m samples
A_2 = factor for calculating limits for the averages chart
$A_2\bar{R}$ = distance of limits from the $\bar{\bar{X}}$ line
$\bar{\bar{X}} \pm A_2\bar{R}$ = limits for the $\bar{X}$ chart (for averages)
$\bar{\bar{X}} \pm I_2\bar{R}$ = limits for individual observations (total dispersion)
D_3, D_4 = conversion factors used in charts for ranges
$D_3\bar{R}, D_4\bar{R}$ = limits for range chart

CHARTS FOR AVERAGES AND RANGES

The $\bar{X}$ or average chart is essentially a simplified method for determining the limits of variation that can be expected in the averages of small samples taken from a constant-cause system, using the range (difference between largest and smallest value) as a measure of dispersion. If a sample average falls outside the estimated control limits, the conclusion can be drawn that almost certainly the out-of-control average point does not come from the same system as the others; in operating terms this indicates a major source of extra variability between the samples, that is, an assignable cause so large that an investigation probably will uncover it and make possible corrective action.

In order to get control of a process—to assure its predictability—it is necessary not only to know its behavior as between rational subgroups (by means of the $\bar{X}$ chart), but also to discover whether or not assignable causes exist *within* the subgroups or samples. For this purpose the range or R chart is used, the limits being set at $D_3\bar{R}$ and $D_4\bar{R}$, whereas the $\bar{X}$-chart limits are $\bar{\bar{X}} \pm A_2\bar{R}$. Thus the two charts reveal the presence of irregularity both between samples ($\bar{X}$ chart) and within samples (R chart). Together they spread a net from which it is difficult for an assignable cause to escape.

Table 2 gives values of A_2, I_2, D_3, and D_4 for sample sizes from $n = 2$ to $n = 10$. Samples larger than 10 are not well adapted to control-chart analysis, because the range is not an efficient measure of dispersion in larger sample sizes.

Circumstances occasionally may require a sample size larger than 10, as when one piece off each of a large battery of machines is taken as a sample, in order to determine whether some factor varying with time is affecting all of the machines similarly; in this type of problem the suspected assignable cause is not between machines but may occur as a function of work flow. Such circumstances are rare, but, when they do occur, and n has to be greater than 10, Technical Appendix A at the end of this chapter can be consulted.

Those who are interested in the mathematical derivation of the symbols used in control-chart work are referred to Technical Appendix B at the end of this chapter.

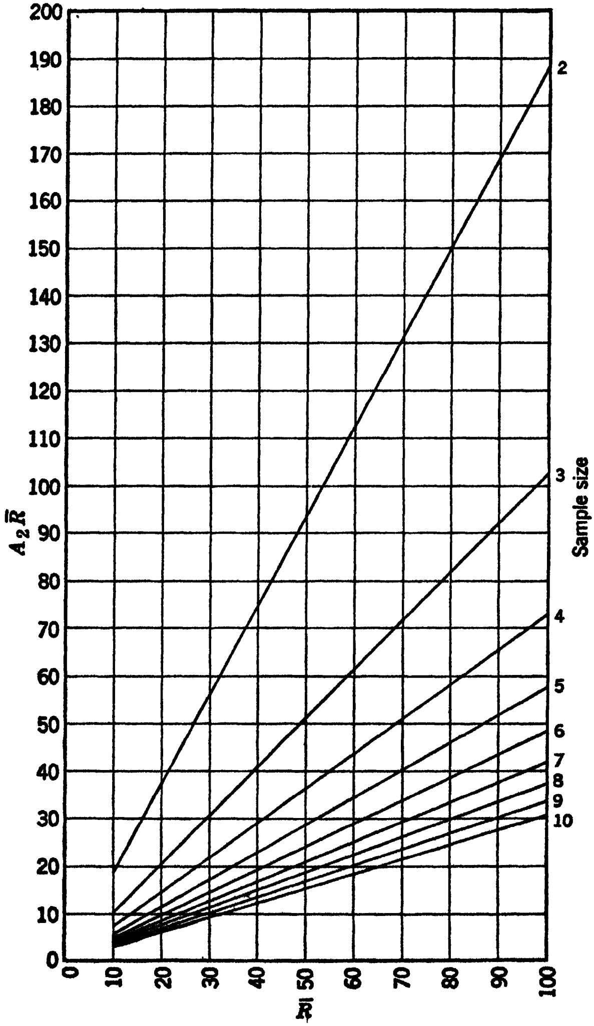


Figure 1. Chart for calculating $A_2\bar{R}$ for sample sizes from $N = 2$ to $N = 10$

For convenience in calculating $A_2\bar{R}$, Figure 1 may be used. Along the bottom of Figure 1 will be found values of $\bar{R}$ from 10 to 100. Along the left-hand margin will be found the values of $A_2\bar{R}$. Each diagonal line represents a sample size from $n = 2$ to $n = 10$. The $\bar{R}$ and $A_2\bar{R}$

scales should be used like a slide rule, by setting the decimal mentally. For instance, the $\bar{R}$ calculated from Table 3 is 0.00516 inch for a sample size of $n = 4$. Set the decimal mentally at 51.6, enter Figure 1 on the bottom scale at 51.6, go up to the diagonal line marked 4, and over to the left-hand margin, reading $A_2\bar{R}$ as approximately 37.5. Resetting the decimal gives $A_2\bar{R}$ a value of 0.00375, accurate to two significant places.

From the data of Table 8, page 60, $\bar{R}$ is found to be 0.02035 inch for $n = 5$. Set the decimal mentally at 20.35, enter Figure 1, read up to the line marked 5, and over to the left-hand margin. $A_2\bar{R}$ is 11.8 or, when the decimal is reset, 0.0118 inch. Actual calculation, by taking A_2 from Table 2, gives (0.577) (0.02035) = 0.01174 inch.

TABLE 2

FACTORS FOR $\bar{X}$ CHART AND R CHART WHEN $n \leq 10$

Number of Observations in Sample (n)	(Limits $= \bar{X} \pm A_2\bar{R}$)			(Limits $= D_3\bar{R}, D_4\bar{R}$)	
	d_2	$A_2 = \frac{3}{d_2\sqrt{n}}$	$I_2 = A_2\sqrt{n}$	D_3	D_4
2	1.128	1.880	2.66	0	3.268
3	1.693	1.023	1.77	0	2.574
4	2.059	0.729	1.46	0	2.282
5	2.326	0.577	1.29	0	2.114
6	2.534	0.483	1.18	0	2.004
7	2.704	0.419	1.11	0.076	1.924
8	2.847	0.373	1.05	0.136	1.864
9	2.970	0.337	1.01	0.184	1.816
10	3.078	0.308	0.98	0.223	1.777

CONTROL-CHART ANALYSIS OF METAL KNOBS

A certain manufacturer received an order for a large quantity of metal knobs with polished edges, for use on the outside of a special type of radio cabinet. Samples that had been turned on a lathe and double-disk-polished on a centerless grinder were submitted to the customer, who approved them. The double disking then was added to the operation sheet, and production was begun.

Double disking is an operation that polishes a round metal object by placing it between two abrasive-covered wheels, set at a fixed dis-

tance apart and rotating in opposite directions. The operation is almost entirely automatic, delivering parts of consistent finish and diameter. Nevertheless it was not long before the foreman complained to the head of the polishing department that the knobs did not "clean

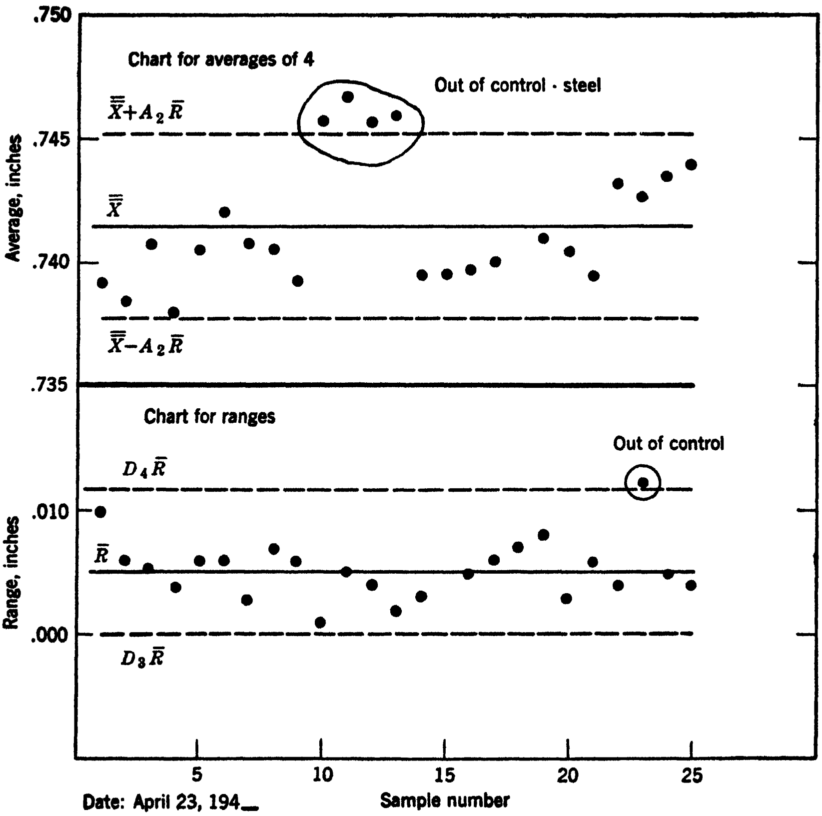


Figure 2a. Control-chart analysis of original data. Diameters of metal knobs

up" properly, and that, if he set the polishing wheels close enough together to give the knobs a satisfactory finish, he burned up his wheels. At the same time a large percentage of the knobs were rejected at final inspection and returned for a second polish. The polishing department head objected to such reworks being charged to his department, claiming that the 0.030-inch tolerance range (0.740 in. ±0.015) at the lathe operation was too loose. In this allegation he was correct. A high finish being required, fine grit had to be used in polishing, and not more than 0.015 inch could be taken off at one pass through the

TABLE 3

CONTROL-CHART ORIGINAL DATA ON METAL KNOBS, IN 1/1,000 INCH

Time	Sample No.	Measurements in Each Sample				Average, $\bar{X}$	Range, R
		a	b	c	d		
Apr 23, 1945							
8:00 am	1	735	745	739	738	739.25	10
8:35	2	742	736	739	737	738.50	6
9:05	3	739	741	739	744	740.75	5
9:25	4	740	736	737	739	738.00	4
10:00	5	738	744	738	742	740.50	6
10:30	6	740	744	739	745	742.00	6
10:55	7	742	739	740	742	740.75	3
11:35	8	741	742	743	736	740.50	7
12:00	9	742	742	736	737	739.25	6
1:00 pm	10	746	746	746	745	745.75 *	1
1:35	11	749	746	748	744	746.75 *	5
2:00	12	745	744	748	746	745.75 *	4
2:30	13	747	745	746	746	746.00 *	2
3:05	14	739	740	741	738	739.50	3
3:30	15	740	739	739	740	739.50	1
3:45	16	742	739	741	737	739.75	5
4:30	17	740	744	738	738	740.00	6
5:00	18	740	740	735	742	739.25	7
5:30	19	745	742	737	740	741.00	8
6:00	20	739	741	742	740	740.50	3
6:30	21	740	740	742	736	739.50	6
6:55	22	744	745	741	743	743.25	4
7:30	23	748	743	744	736	742.75	12 *
8:00	24	741	745	742	746	743.50	5
8:35	25	743	745	746	742	744.00	4
					Totals	18,536.25	129
					$\bar{\bar{X}} =$	741.45	
					$\bar{R} =$		5.16

$\bar{\bar{X}} = 741.45$
$\bar{R} = 5.16$
A_2 for samples of 4 (Table 2) $= 0.729$
D_3 for samples of 4 (Table 2) $= 0$
D_4 for samples of 4 (Table 2) $= 2.282$
$A_2\bar{R} = 0.729 \cdot 5.16 = 3.76$
$\bar{X}$ limits $= \bar{\bar{X}} \pm A_2\bar{R} = 741.45 \pm 3.76 = 745.21, 737.69$
R limits $= D_3\bar{R} = 0 \cdot 5.16 = 0$
$D_4\bar{R} = 2.282 \cdot 5.16 = 11.76$

* Out of control.

grinder. As long as the knobs were not polished, the 0.030-inch tolerance range was all right; but with the new specification as to the appearance of the knobs, it took two passes through the machine, involving two setups, to polish all the knobs properly. In order to avoid the double work, an investigation of the turning operation at the lathe was undertaken, and proved to be such an excellent illustration of the control-chart technique that it is described here in some detail.

Four successive knobs from the lathe were measured with a micrometer and the measurements recorded every half-hour during two shifts' production. The data are shown in Table 3 and in Figure 2a. Calculations are shown at the bottom of Table 3. Each point on the chart for averages represents the arithmetic mean of a sample of four successive observations. The corresponding point on the range chart represents the range (largest minus smallest value) in that sample. A group of four points—10, 11, 12, and 13—were out of control on the average chart and 23 was out on the range chart.

Inquiries were made about the cause of the four out-of-control sample averages. It developed that after lunch on April 23, work started on a new bar of steel. By 3 o'clock it was used up. Tracing down this clue revealed that the bar came from a lot of steel that "Rockwelled" slightly higher than the general run of steel purchased, and that all the extra hard bars came from one supplier.

An assignable cause having been discovered, the four offending samples were eliminated from the data, and limits were recalculated as shown in Figure 2b and below:

$$\bar{\bar{X}} = 0.74001 \text{ inch (21 sample averages)}$$
$$\bar{R} = 0.00552 \text{ inch (21 sample averages)}$$
$$A_2 = 0.729$$
$$D_3 = 0$$
$$D_4 = 2.282$$
$$A_2\bar{R} = 0.729 \cdot 0.00552 = 0.00402 \text{ inch}$$
$$\bar{X} \text{ limits} = \bar{\bar{X}} \pm A_2\bar{R} = 0.74001 \pm 0.00402 = 0.74403,\ 0.73599 \text{ inch}$$
$$R \text{ limits: } D_3\bar{R} = 0$$
$$D_4\bar{R} = 2.282 \cdot 0.00552 = 0.01260$$

All points on the average chart were now in control. No. 23 which showed an out-of-control range on the first calculation, was in control on the second. Figure 2b shows a controlled process, with no single cause of variation worth looking for.

Although the process gave no evidence of lack of control, it was not satisfactory. Dispersion was still too great. Of the 84 individual measurements in Table 3 (excluding the out-of-control samples), the lowest was 0.735 inch, the largest 0.748 inch, a range of 0.013 inch.

This was the dispersion of the *controlled* process. Since it was known that differences in raw materials and perhaps other factors previously had thrown the process out of control, the dispersion in the past had been considerably greater—probably close to the 0.030 inch permitted by the lathe operation tolerances.

It is an axiom in statistical-quality-control work that, once a process has reached and maintained a state of statistical control, a further

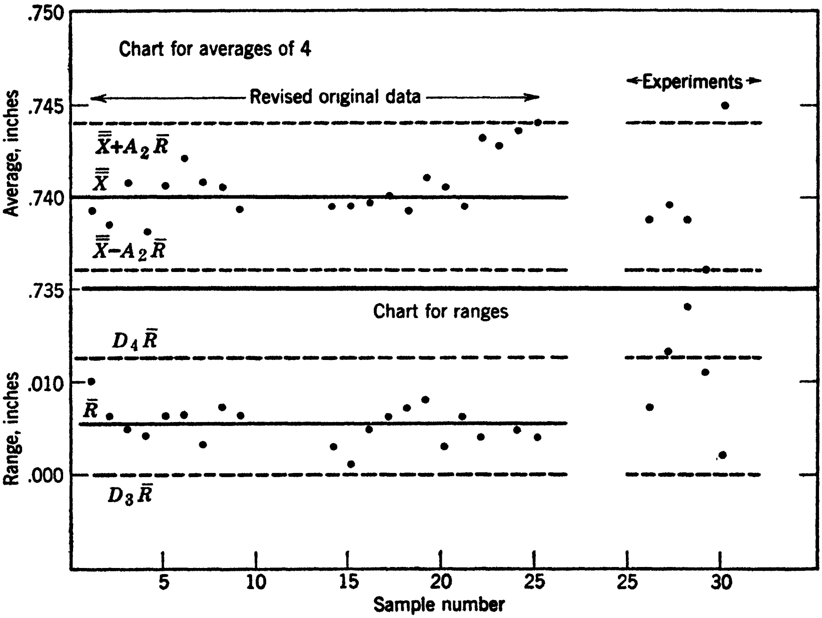


Figure 2b. Revised analysis of original data. Diameter of metal knobs

reduction in dispersion cannot be gained by any temporary or partial expedients; further improvement comes only from some basic change in the manufacturing procedure. On the morning of April 24 experiments were begun, aimed at cutting the variation in half. Results of these attempts are shown in Figure 2b and in Table 4.

None of the trials with various types of tooling or different methods of setup, or even with the most skillful operators were successful until a re-examination of the original data showed that all four of the out-of-control points (samples 10–13) had ranges less than the average range value of $\bar{R} = 0.00516$. As the final experiment of the day, therefore, a bar of the steel that previously had been a cause of irregularity was used. The control-chart sample taken at 4 o'clock showed an average of 0.745 inch and a range of only 0.002 inch. The rest of the day was

spent testing a number of the bars to determine whether that particular kind of steel varied less in hardness than other kinds. It did seem to be

TABLE 4

EXPERIMENTS ON METAL-KNOB PROCESS

Time	Sample No.	Measurements in Each Sample				Average, $\bar{X}$	Range, R
		a	b	c	d		
Apr 24, 1945							
9:00 am	26	742	744	742	737	738.75	7
11:00	27	746	736	743	733	739.50	13
1:30 pm	28	733	735	751	746	738.75	18
3:30	29	739	741	734	730	736.00	11
4:00	30	746	744	746	744	745.00	2

more consistent, and so the next morning the turning operation was set up and carried on throughout the day using only the steel from that source. Table 5 and Figure 2c show the results.

The last step indicated that the process was in control with a mean, $\bar{X}$, of 0.74620 inch and an average range, $\bar{R}$, of 0.00423 inch. Control limits for the $\bar{X}$ chart were 0.74928 and 0.74312 inch. Control limits for individual pieces were found by taking $\bar{X} \pm A_2\bar{R}\sqrt{n} = \bar{X} \pm I_2\bar{R}$:

$$\bar{X} \pm I_2\bar{R} = 0.74620 \pm 1.46 \cdot 0.00423 = 0.74620 \pm 0.00618$$
$$= 0.752,\ 0.740 \text{ inch}$$

Specifications were $0.740 \pm 0.015 = 0.755$, 0.725 inch. As long as statistical control was maintained, the upper individual limit of the *process*, at 0.752 inch would fall below the upper *tolerance* limit of 0.755 inch, and the total variation would be not more than $0.752 - 0.740 = 0.012$ inch. This met the requirement of the double-disking operation where a total variation up to 0.015 inch could be taken at a single cut. On the new basis, therefore, the process was in control *at a satisfactory level.*

In achieving this result, the following steps so far had been taken:

1. Collection of original data (samples 1–25).
2. Calculation and plotting of control charts for $\bar{X}$ and R.
3. Engineering investigation of out-of-control points.

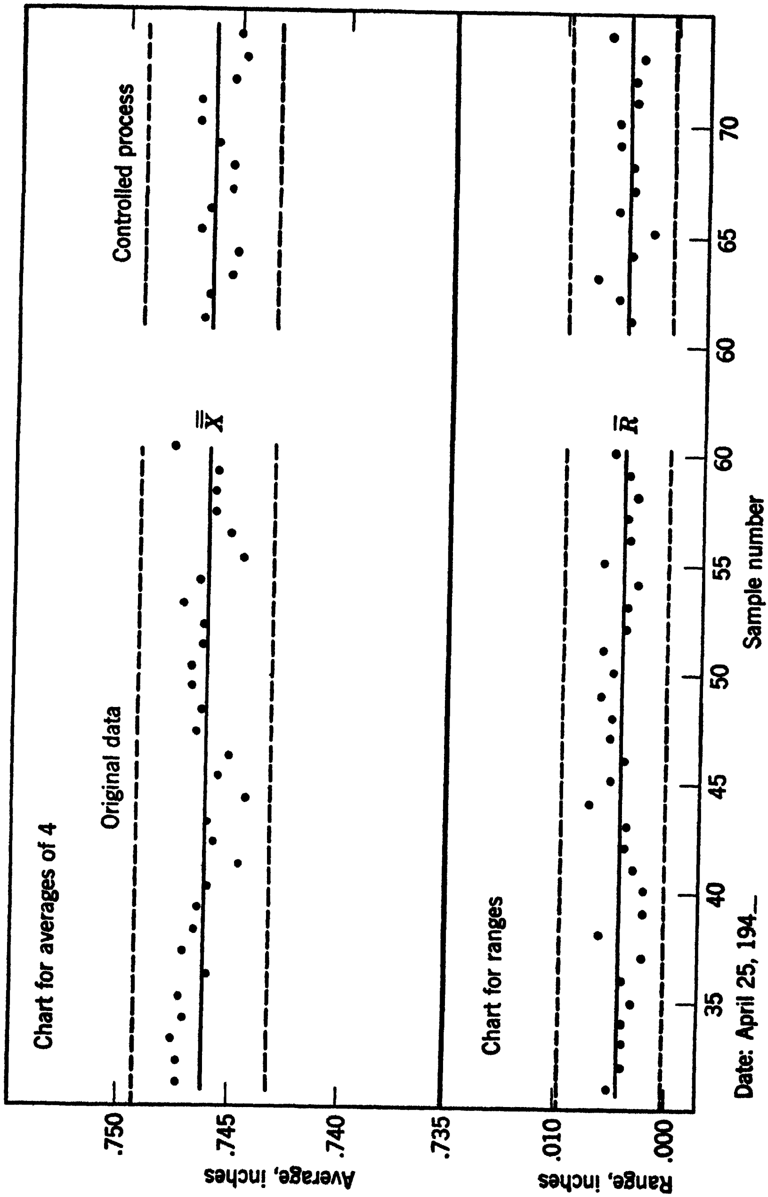


Figure 2c. Control-chart analysis of improved process. Diameter of metal knobs

TABLE 5

CONTROL-CHART DATA ON IMPROVED-PROCESS METAL KNOBS, IN 1/1,000 INCH

Time	Sample No.	Measurements in Each Sample				Average, $\bar{X}$	Range, R
		a	b	c	d		
Apr 25, 1945							
8:00 am	31	749	744	749	747	747.25	5
8:30	32	745	749	746	749	747.25	4
9:00	33	745	749	749	747	747.50	4
9:30	34	746	745	748	749	747.00	4
10:00	35	745	748	748	748	747.25	3
10:35	36	744	748	747	745	746.00	4
10:55	37	746	748	748	746	747.00	2
11:30	38	745	743	749	749	746.50	6
12:00	39	747	745	747	746	746.25	2
1:05 pm	40	745	747	747	745	746.00	2
1:30	41	744	746	743	745	744.50	3
2:00	42	748	744	746	745	745.75	4
2:35	43	747	745	748	744	746.00	4
3:00	44	748	741	744	744	744.25	7
3:30	45	747	746	747	742	745.50	5
4:00	46	744	744	748	744	745.00	4
4:25	47	749	744	747	746	746.50	5
5:00	48	748	747	747	743	746.25	5
5:30	49	749	749	746	743	746.75	6
5:55	50	749	746	748	744	746.75	5
6:25	51	749	748	743	745	746.25	6
7:00	52	748	744	747	746	746.25	4
7:30	53	748	745	749	747	747.25	4
8:30	54	748	746	745	747	746.50	3
9:00	55	742	745	748	743	744.50	6
9:30	56	746	747	744	743	745.00	4
10:05	57	746	748	744	745	745.75	4
10:30	58	746	746	747	744	745.75	3
11:00	59	745	746	748	744	745.75	4
11:25	60	749	749	749	744	747.75	5
					Totals	22,386.00	27
					$\bar{\bar{X}} =$	746.20	
					$\bar{R} =$		4.23

$A_2 = 0.729$

$D_3 = 0$

$D_4 = 2.282$

$A_2\bar{R} = 3.08$

$\bar{\bar{X}} \pm A_2\bar{R} = 746.20 \pm 3.08 = 749.28,\ 743.12$

$D_3\bar{R} = 0$

$D_4\bar{R} = 2.282 \cdot 3.08 = 9.65$

4. Recalculation and plotting of original data, indicating a state of statistical control but at an unsatisfactory level.

5. Further engineering studies to improve the process basically (samples 26–30).

6. Collection of new original data on the improved process (samples 31–60).

7. Calculation and plotting of new data on control charts.

Then the last and most important step was taken: the control limits, being satisfactory, were accepted as process limits *for future production,* that is, a prediction was made that, as long as the control limits were held, there would be little if any trouble at double disking. A continuous control-chart record was set up for the turning operation, beginning on the morning of April 26, with inspection at hourly

TABLE 6

CONTROL-CHART DATA PLOTTED AGAINST STANDARD VALUES FOR DIAMETERS OF KNOBS

Time	Sample Measurement in Each Sample (0.001″)					Average, $\bar{X}$	Range, R
	No.	*a*	*b*	*c*	*d*		
Apr 26, 1945							
7:00 am	61	746	749	745	746	746.50	4
8:00	62	748	748	746	743	746.25	5
9:00	63	746	744	749	742	745.25	7
10:00	64	745	745	747	743	745.00	4
11:00	65	746	746	748	747	746.75	2
12:00	66	748	748	743	746	746.25	5
1:30 pm	67	744	748	745	744	745.25	4
2:30	68	748	745	744	744	745.25	4
3:30	69	744	749	747	744	746.00	5
4:30	70	746	749	748	744	746.75	5
5:30	71	746	747	749	745	746.75	4
6:30	72	744	747	747	743	745.25	4
7:30	73	746	746	744	743	744.75	3
8:30	74	749	749	743	749	747.50	6

Standard Values: $\bar{X}' = 0.74620$ inch
$R' = 0.00423$ inch
$\bar{X}'$ limits: 0.74928, 0.74312 inch
R' limits: 0.00965, 0 inch
Individual limits: 0.752, 0.740 inch

instead of half-hourly intervals.[1] The $\bar{\bar{X}}$ and limit lines on the chart for averages, and the $\bar{R}$ and $D_4\bar{R}$ lines on the chart for ranges were extended into the future before production was started. In the terminology of statistical control, $\bar{\bar{X}}$, the process mean, became $\bar{X}'$, the aimed-at standard mean; $\bar{R}$ became R', the standard average range; and the control-chart limits became the standard limits. As production went forward, the inspection data were gathered and analyzed as before; each sample's average and range were plotted against the *projected* limit lines in order that a statistically controlled process might be maintained with the agreed-upon average and dispersion. Unless this were done, there would be no guarantee that the process would remain satisfactory, no assurance that the potential benefits of the engineering–statistical analyses actually would be realized. Control for the future is the goal and end result of all the preceding research.

Table 6 and the right-hand portion of Figure 2c show a few of the observations that were plotted against the previously extended control limits.

WHEN IS A CONTROL CHART VALID?

So powerful an analytical tool is the control chart and so easy to use that it is recommended as an approach to methods engineering wherever inspection can be carried out by taking small homogeneous rational subgroups at frequent intervals. It is important to know, however, whether the subgroups have been chosen correctly. If the subgroups do not succeed in isolating the real cause of trouble, the control chart may fail in its function and throw discredit on the whole technique. Probably three out of four first attempts at control charts are unsuccessful because of inexperience and the failure to apply a scientific criterion of appropriateness.

There is no substitute for experience and common sense in making the most out of control-chart analysis. If the chart *looks* right it is apt to be right. In general, a chart looks right if:

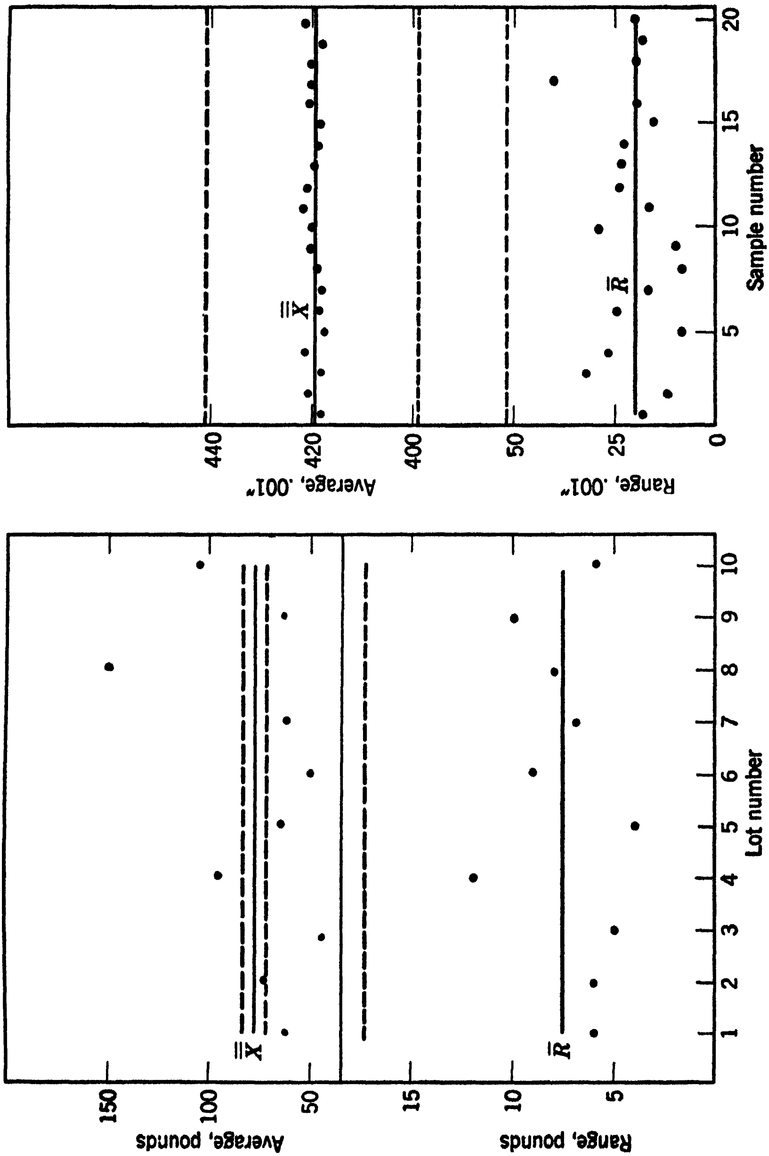
(*a*) For a process in control, the $\bar{X}$ points spread out from the mean at least halfway to the control limits on each side.

(*b*) For a process not in control, the limits cover at least half of the total spread of $\bar{X}$ points.

These criteria are rough and ready ones, derived from the author's experience; more refined methods are required if any real doubt exists as to the validity of the chart.

Figure 3 shows two extreme cases. In Figure 3a the control limits

[1] Such reduced inspection is one of the major advantages of knowing that an operation is in a state of *statistical* control.



(b) Wall thickness of metal tubing

(a) Breaking strength of cotton string

Figure 3. Nonvalid con

look too close together for the data. This indicates the presence of assignable causes so great that it may be difficult if not impossible to achieve statistical control. The common denominator that should exist for all the groups seems to be lacking. There may, in such a case, be no possible way in which control can be obtained with this particular breakdown; and, if the control chart cannot be used to help attain a state of control, its usefulness is greatly circumscribed.

In Figure 3b, the limits are too wide apart; the data cluster too closely around the $\bar{\bar{X}}$ line. This kind of picture can be interpreted to mean that the variation *within* the samples, upon which the calculation of control limits depends, is much larger than the variation *between* samples. If an assignable cause exists, it probably will be found inside the samples; that is, the subgrouping has not succeeded in its objective of isolating a real cause of trouble. Some other subgrouping then should be tried.

Limits Too Close Together

In Figure 2a there were ten lots of string, each point on the chart representing four tests of tensile strength from one of the lots. Table 7 shows the data from which Figure 3a was plotted.

TABLE 7

BREAKING STRENGTH OF COTTON STRING FROM 10 DIFFERENT LOTS

Lot No.	Test Values on Each Lot, Lb				Average, $\bar{X}$	Range, R
	a	*b*	*c*	*d*		
1	63	66	62	60	62.75	6
2	72	78	74	72	74.00	6
3	47	42	46	44	44.75	5
4	97	88	100	100	96.25	12
5	63	67	66	66	65.50	4
6	55	46	50	52	50.75	9
7	60	59	66	65	62.50	7
8	148	146	154	149	149.25	8
9	63	64	68	58	63.25	10
10	106	104	110	104	106.00	6

$\bar{\bar{X}} = 77.50$
$\bar{R} = \quad 7.6$

Since the purchaser had, in this receiving inspection, no control over the manufacturing processes of his various suppliers, it was obvi-

ous that though assignable causes or significant differences existed among his sources, there was nothing he could do about them. Since he could not take action to improve the quality of his material, the only thing he could do would be to choose the most satisfactory supplier and order only from the one source. If that were not practical or economical, he could put more rigid receiving inspection on the poor-quality vendors. If any one vendor, though possibly out of control, met the purchaser's specifications, he could continue to accept material from that source, with receiving inspection sufficient to prevent acceptance of substandard lots. In any case, the chart as originally set up should not be continued. It might be preferable to keep a separate control chart on each supplier; but the chart as shown in Figure 3a has little value, because the assignable causes, though known, cannot be corrected.

Occasionally such a situation arises in an integrated factory where, for instance, some ersatz or substitute material has to be used, or a workman, though admittedly unsatisfactory, has to be kept on the job because he cannot be replaced. The quality-control practitioner who continues to insist upon stopping production and investigating assignable causes, when they are already known and cannot be eliminated, creates misunderstanding and conflict between himself and the factory workers and foremen.

Another pitfall of the same sort, though existing for a different reason, is illustrated in Chapter 5, Case History VII. Chart VIIa looks very much like Figure 3a, but the similarity is only superficial. Case History VII presents a problem in subgrouping, a type of problem frequently met in control-chart practice, and one that, if unsolved, can only embroil the quality-control man in the danger of continually pointing out assignable causes of trouble that cannot be found.

Limits Too Far Apart

Figure 3b illustrates a less frequent kind of invalid control chart, when the points cluster too closely about the mean and too far from the control limits. Here the danger lies in not finding an assignable cause when one exists.

Figure 3b illustrates a process of drawing out seamless metal tubing where difficulty was being encountered because the wall thickness was not consistent. At times it would seem to be too thick, and at other times too thin. Although tolerances were being held, trouble was being met in subsequent operations because of the variability in wall thickness. Inspection of wall thickness customarily was per-

formed by measuring the thickness at three points on one end of the tube, the points being 120 degrees apart, the starting point being chosen at random. Twenty samples, each comprising three measurements, are shown in Table 8, and the averages and ranges are plotted in Figure 3b.

TABLE 8

MEASUREMENTS OF WALL THICKNESS OF METAL TUBING (IN 1/1,000 INCH)

Sample No.	Measurements 120 Degrees Apart			$\bar{X}$	R
	a	*b*	*c*		
1	414	412	430	418.7	18
2	428	418	416	420.7	12
3	432	426	400	419.3	32
4	437	410	415	420.7	27
5	419	413	422	418.0	9
6	432	417	407	418.7	25
7	409	426	420	418.3	17
8	416	418	424	419.3	8
9	418	426	416	420.0	10
10	415	437	408	420.0	29
11	416	433	416	421.7	17
12	418	410	434	420.7	24
13	412	435	412	419.7	23
14	432	409	416	419.0	23
15	425	421	410	418.7	15
16	435	415	420	420.0	20
17	400	440	420	420.0	40
18	430	420	410	420.0	20
19	428	410	417	418.3	18
20	412	419	432	421.0	20

$\bar{\bar{X}} = 419.64$

$\bar{R} = 20.35$

A review of the figures revealed a tendency for one of the three measurements to be very much larger or smaller than the other two. An investigation revealed that, at the first stage of the process, when the tube was bored from the billet and before it was drawn out at succeeding operations, the metal became heated and tended to flow slightly. Following the law of gravity, the portion of the tube at the bottom of the mandril became thicker than elsewhere, the thickening

effect covering an area from one third to one half of the circumference. At inspection of subsequent drawing operations, sometimes one measurement, sometimes two, would hit the thickened part of the tube, making the sample ranges excessively large.

The unevenness of the wall thickness was reduced by the tube being rotated at intervals during the boring operation, thus equalizing the effect of metal flow. Thereafter no further difficulty from this cause was encountered, and control charts kept of wall thickness at various stages of the process became effective in improving the quality of the finished tubing.

ECONOMICAL CONTROL LIMITS

Problems such as those previously discussed point out the two kinds of mistakes, both of them uneconomical and costly, that can be made in quality-control work:

1. Investigating assignable causes when they do not exist or cannot be corrected. Investigations usually require the services of a competent engineer; they call for extra inspection and analytical work; they demand special attention from supervisors, foremen, and workers. Altogether, the cost may be considerable. If it is incurred without a strong probability of finding the cause of excessive variation, the purpose of statistical quality control is vitiated, and the technique loses acceptance among those who should benefit from it and use it most.

2. Not investigating when assignable causes do exist. This kind of error, though less frequently made, may be even more costly than the first kind, because it may perpetuate inefficiencies and high costs that otherwise could be avoided. Men who have had experience with the results obtained from control charts probably will agree that the danger of an occasional fruitless investigation is preferable to the danger of not finding an assignable cause when it exists.

There is nothing sacred or mystical about the control limits generally used in quality-control work. When a beginner starts to study the subject, he may ask, "Why are the factors A_2, D_3, and D_4 set at the values given in the tables? Why not twice as large, or two thirds as large, or one half as large?" The answer is somewhat as follows:

The limits are set so that, in a process giving evidence of a state of statistical control, as indicated by a properly designed control chart, only about three points in every 1,000 will fall outside the limits by chance alone. That is, about 0.3 per cent of the time an assignable cause will be indicated when none exists; or, put another way, the probability is approximately 0.997 that an assignable cause exists

when the chart indicates lack of control. This has been found, from many years of experience in many plants, to strike an economic balance between the two kinds of errors: looking for an assignable cause when it does not exist, and not looking for it when it does exist.

The author's experience has borne out the experience of others. It is often difficult to achieve statistical control, even with the generous limits established by common practice. Occasionally, narrower limits may be advisable, using $\frac{2}{3} A_2$ or $\frac{1}{2} A_2$, in the formulas. But before such modified limits are accepted, the rational subgrouping, the inspection techniques, and the characteristics of the process should be studied exhaustively. Modified limits should be used only as a last resort.

Wider limits than those accepted by common usage—that is, A_2, D_3, and D_4 multiplied by a factor larger than 1—are not advisable. The author has never come across a case in factory processes where wider limits did anything except hide findable causes of excessive variation.

SAMPLE SIZE

In using control charts for variables—$\bar{X}$ and R charts—the most important decision, after determining what operation to analyze, is to select what rational subgrouping to use. When that decision has been made, the next question usually is what sample size to use. Often a characteristic of the machine or the process makes the choice of sample size easy. More often other considerations dictate sample size. Some of them are:

Sample size of 2. Should not be used unless there is a good engineering reason for it. The estimates of limits for $n = 2$ are apt to be erratic. Remember that control limits are merely *estimates.* Estimates based on samples of 2 contain a larger error than those based on larger samples.

Sample size of 3. Can be used satisfactorily, but has the drawback of being somewhat difficult to calculate, since recurring decimals appear in the $\bar{X}$ column. If it is difficult or expensive to collect the original data or to make the measurements, samples of $n = 3$ may be used to reduce the amount of work required in getting an adequate number of samples, m (see next section).

Sample size of 4. A very good sample size to use, stable yet sensitive. Fours are advisable if measurements are made in quarters of an inch, to the nearest quarter-thousandth, and so on.

Sample size of 5. Also very satisfactory. Best to use if measurements are in decimals. Most people find 5 easier to divide by in calculating averages than any other number, though sometimes 4 is pre-

ferred. The drawback when $n = 5$ is that relatively more measurements must be made to get a minimum number of samples, m.

Sample size of 6. Not recommended unless engineering or subgrouping reasons call for it. Two samples of three are generally better than one sample of six.

Sample sizes larger than 6. Not recommended. Use 3, 4, or 5 instead. If $n > 10$ the range is not an efficient measure of dispersion, and Technical Appendix A should be consulted.

NUMBER OF SAMPLES

The number of samples of size n is designated by m. For analyzing past data as few as ten samples are permissible. After the preliminary analyses and investigations have been completed and the chart is ready to go as a control for production, it is wise to have *more than 20 samples* before projecting $\bar{\bar{X}}$ and limit lines into the future. The author uses at least 30 and if possible as many as 50 samples in calculating $\bar{\bar{X}} \pm A_2\bar{R}$, $D_3\bar{R}$, and $D_4\bar{R}$ for guiding future production.

Sampling errors vary inversely as the square root of the number of observations. Holding for large samples as well as small ones, this means that, in a sample 4 times as large as a selected one, the error of estimate of the average will be one half as great; for a sample 100 times as large the error will be one tenth as great. Before control-chart limits are extended as guides for the future, they should contain as small a sampling error as is economically possible. The size of that error depends upon the total number of observations, N, used in calculating the limits. During the experimental period, while assignable causes are being hunted down, as few measurements as seem necessary may be made, but, once the chart is set for predicting the future, at least 100 observations should be used in calculating the mean and the limits. Then, if the sample size is 3, the number of samples m should be at least 34; if $n = 4$, m should be at least 25; if $n = 5$, m should be at least 20.

The dangers of single small samples often are not realized by people unacquainted with statistics. In factory inspection, a single sample of two or three observations usually will include a sampling error so large as to make the information unreliable; in some circumstances even 100 measurements will not give the desired reliability. The finest engineering research cannot reduce the chance errors due to sampling; only an adequate sample can do that. It is imperative, therefore, in the engineering-statistical teamwork required by effective control charts, that an adequate total sample, N, be collected and analyzed before control limits are projected for the guidance of future production.

RUNS

A recent development in control-chart technique is the theory of runs. A run is a movement of successive points in the same direction. There are two kinds of runs:

Runs up or down. A series of points, each higher than the preceding one, is a run up; a series of points, each lower than the preceding one, is a run down.

Runs above or below the average, are, by a similar definition, a series of successive points all falling above or below the average.

Runs usually indicate the presence of incipient or developing assignable causes. They frequently give advance warning of a departure from control standards. For both types of runs the author has found the following criterion for action useful:

Five successive points: Be on guard for future developments.

Six successive points: Start investigation.

Seven successive points: Take action.

The different kinds of action to be taken are:

Runs up or down.

(*a*) May indicate approaching lack of control. This is a definite danger signal that, if heeded, often can prevent bad work before it appears. Sometimes the progression (rising or falling) is so rapid that the third or fourth point may fall outside control limits. In such a case, action should be taken when any point in the run approaches close to either control limit.

(*b*) May indicate a permanent feature of the process, such as tool wear, which appears as a pattern of recurring runs on the chart. Treatment should be as in Chapter 5, Case History VII.

Runs above or below average.

(*a*) May indicate a shift in the average, tending to produce out-of-control points. Investigate and correct it, perhaps by recalculating the average and the limits.

(*b*) May indicate a break or lack of continuity in the process due to a change in raw materials, resharpening or changing of tools, change of operator or shift, and the like. An investigation is called for, with action dependent upon the findings.

WHEN TO RECALCULATE LIMITS

A control chart, once set up and projected into the future, needs to be reviewed at intervals in order to keep it abreast of current developments. As changes and improvements take place, the control chart

should reflect them. It is this continual revaluation of the process that brings about much of the *progressive* improvement achieved through control charts. Standards should not be fixed and final, but should lead to increasing efficiency step by step, keeping the demands upon the process within the limits of which it is capable, but gradually increasing these demands as the process is improved. One of the secrets of progress is to set an advancing goal which nevertheless is at any time possible of attainment. Revaluing of the control chart provides a technique for doing this with factory processes.

Suppose that a control chart has been set up and projected for future guidance. After a number of samples have been taken and plotted against the projected limits, one of two things has happened: either the chart is still in control, or it evidences lack of control. The revaluing procedure is different in each case.

1. If the chart remains in control. If all the $\bar{X}$ points continue to fall inside the projected limits and no runs above or below average appear either on the $\bar{X}$ chart or on the R chart, probably no change has occurred in the process. In that case, allow approximately 500 *individual* observations to accumulate. Recalculate $\bar{\bar{X}}$, $A_2\bar{R}$, $D_3\bar{R}$, and $D_4\bar{R}$ from the past 500 observations, and extend the new limits to cover the next 500.

2. If the chart shows one or more points out of control. If the out-of-control points are occasional, suggesting unpredictable in-and-out assignable causes, limits should be recalculated for every 100 individual observations while efforts to find and eliminate the trouble continue.

If there are runs up-or-down, trends due to tool wear or other constant causes may be present and should be hunted for. While the search is going forward, recalculate averages and limits every 100 observations.

If runs above or below average appear on the $\bar{X}$ or R charts, a shift in the value of $\bar{\bar{X}}$ or $\bar{R}$ is indicated. The chart should be recalculated as soon as seven successive points occur in a run, with the preceding 100 observations being used as a basis.

COMMON SENSE AND STATISTICS

Common sense should always supplement these rules. Liberal use of practical good judgment is more effective than a multitude of statistical tests. When something looks wrong, either in the figures or on the chart, a little clear thinking and the application of lessons learned from experience will help to unearth the difficulty. Entire reliance upon statistics, especially the more elaborate formulas, is the hall-

mark of an inexperienced statistician. Common sense, good judgment, and a thorough practical knowledge of the process are absolutely essential to successful control-chart work. Statistical tests can best be used where real doubt exists, where decisions may be expensive, or where basic research for determining permanent specifications and quality standards is being carried out.

Nowhere is the combination of experience and theory more necessary than in engineering investigations. These are an integral part of the control-chart technique; its value is measured largely by the success attending its use. Knowing what to hunt for, how to carry on the investigation, and when to stop is difficult. No rules can be laid down for weighing in advance the cost against the value of special process research. It may depend not only upon lack of control but also upon the relationship of the control limits to the tolerance limits and to other quality features such as finish or shape. The amount of money that can be saved by gaining control, set against the cost of achieving it, is, in the last analysis, the economic criterion by which the value of control charts will be judged. No matter how interesting they may be, no matter how impressive or how spectacular, only if they pay their own way in dollars and cents on the floor of the factory will control charts accomplish their purpose.

TECHNICAL APPENDIX A

When the sample size must be larger than 10, the chart for standard deviations (σ) is used, with control limits set at $\bar{\bar{X}} \pm A_1\bar{\sigma}$, where

$$A_1 = \frac{3}{\sqrt{n}} \cdot \frac{1}{c_2}$$

and c_2 is a factor for converting the average standard deviation of small samples ($\bar{\sigma}$) into the estimated standard deviation of the universe. Values of A_1 and c_2 for sample sizes from $n = 11$ to $n = 25$ are given in Table 9.

TABLE 9

FACTORS FOR $\bar{X}$ CHART WHEN SAMPLES ARE LARGER THAN $n = 10$

(LIMITS—$\bar{\bar{X}} \pm A_1\bar{\sigma}$)

Number of Observations in Sample (n)	c_2	A_1
11	0.9300	0.973
12	0.9359	0.925
13	0.9410	0.884
14	0.9453	0.848
15	0.9490	0.817
16	0.9523	0.788
17	0.9551	0.762
18	0.9577	0.738
19	0.9599	0.717
20	0.9619	0.698
21	0.9638	0.680
22	0.9655	0.662
23	0.9670	0.647
24	0.9684	0.632
25	0.9697	0.619
26 or more	$\sqrt{\frac{n-1}{n}}$	$\frac{3}{\sqrt{n-1}}$

TECHNICAL APPENDIX B

The factors d_2, A_2, I_2, D_3, and D_4 are used in estimating the values of $t\sigma$ and $t\sigma_{\bar{X}}$ in a statistically controlled cause system, where there is no significant difference between samples or within samples.

$$A_2 = \frac{3}{\sqrt{n}} \cdot \frac{1}{d_2}$$

$$A_2\bar{R} = \frac{3\bar{R}}{d_2\sqrt{n}} \quad \text{for samples of size } n$$

$$\bar{\bar{X}} \pm A_2\bar{R} = \text{limits for } \bar{X} \text{ chart}$$

$$I_2 = A_2\sqrt{n} = \frac{3}{d_2} = \text{limits for individual values}$$

$$D_3, D_4 = \text{conversion factors used in charts for ranges.}$$

D_3 and D_4 are derived as follows:

R chart limits are set at $\bar{R}_n \pm 3\sigma_R$ for samples of size n.

When samples of size n show an $\bar{R}$ with the value of $\bar{R}_n$, the standard deviation of the (assumed) normal universe from which the samples came can be estimated from the equation,

$$\sigma' = \frac{\bar{R}_n}{d_2}, \quad \text{or} \quad \bar{R}_n = d_2\sigma' \qquad \text{and the limits become}$$

$$d_2\sigma' \pm 3\sigma_R$$

Dividing by σ' in order to get the limits in terms of the universe standard deviation, they become

$$d_2 \pm \frac{3\sigma_R}{\sigma'}$$

Factoring for d_2, the limits in terms of σ' become

$$\frac{1}{d_2}\left(1 \pm \frac{3\sigma_R}{d_2\sigma'}\right)$$

Values of $\frac{\sigma_R}{\sigma'}$, that is, the relationship of the standard deviation of sample ranges to the standard deviation of the normal parent distribution, have been determined by L. H. C. Tippett and Egon S. Pearson.[2] In Table 2 the values of D_3 and D_4 are determined as follows:

$$D_3 = \frac{1}{d_2}\left(1 - \frac{3\sigma_R}{d_2\sigma'}\right)$$

If D_3 is negative, the lower limit is taken as zero.

$$D_4 = \frac{1}{d_2}\left(1 + \frac{3\sigma_R}{d_2\sigma'}\right)$$

As Shewhart[3] points out, the distribution of ranges of small samples is closely connected with the functional form of the parent universe. If the universe is not normal, the range chart is apt to be inaccurate. In practice, however, this reservation as to the use of ranges rarely has a significant effect.

[2] L. H. C. Tippett, *Biometrika*, Vol. XVII, December 1925, pp. 364–87; Egon S. Pearson, *Biometrika*, Vol. XXIV, November 1932, pp. 404–07.

[3] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*, pp. 203–05.

CHAPTER 4

p AND pn CHARTS

Inspection by variables, as required for $\bar{X}$ and R charts, is sometimes uneconomical or even impossible. An automatic turning, for instance, may have a dozen dimensions and therefore may call for a dozen different $\bar{X}$ and R charts; but with a go–no-go gauge, a record of over-all quality can be kept by simply noting the percentage of pieces that fail to gauge. Hot forgings have defects such as pits, laps, and cold shuts that are not measurable at all: they can be classified only as passable or not passable. Sometimes a product either succeeds or fails a test: certain types of detonators fall into this category. Frequently at the final inspection of a product there may be so many possible causes for rejection, both in dimensions and in finish, that a separate record of each cause is out of the question. Under such circumstances the best procedure may be to record simply the number of tests performed (by gauging or otherwise) and the number of failures. This is known as inspection by attributes, that is, by classification of the product as either good or bad, as either acceptable or not acceptable. For analyzing and controlling quality on this basis, p and pn charts can be used.

p CHARTS—CONSTANT SAMPLE SIZE

p charts are constructed by recording the results of at least 20 successive inspections, calculating the percentage of defectives found in each, taking the weighted average of all the percentages, and plotting the average or $\bar{p}$ and the limits as found from the formula,

$$\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where n is the number of tests performed at each inspection and $\bar{p}$ is the average weighted fraction defective. The author has found it usually more convenient to express $\bar{p}$ as a per cent rather than as a decimal fraction; for example, as 5 per cent instead of 0.05. Then the formula becomes,

$$\bar{p}\% \pm 3\sqrt{\frac{\bar{p}\%(100-\bar{p}\%)}{n}}\ \%$$

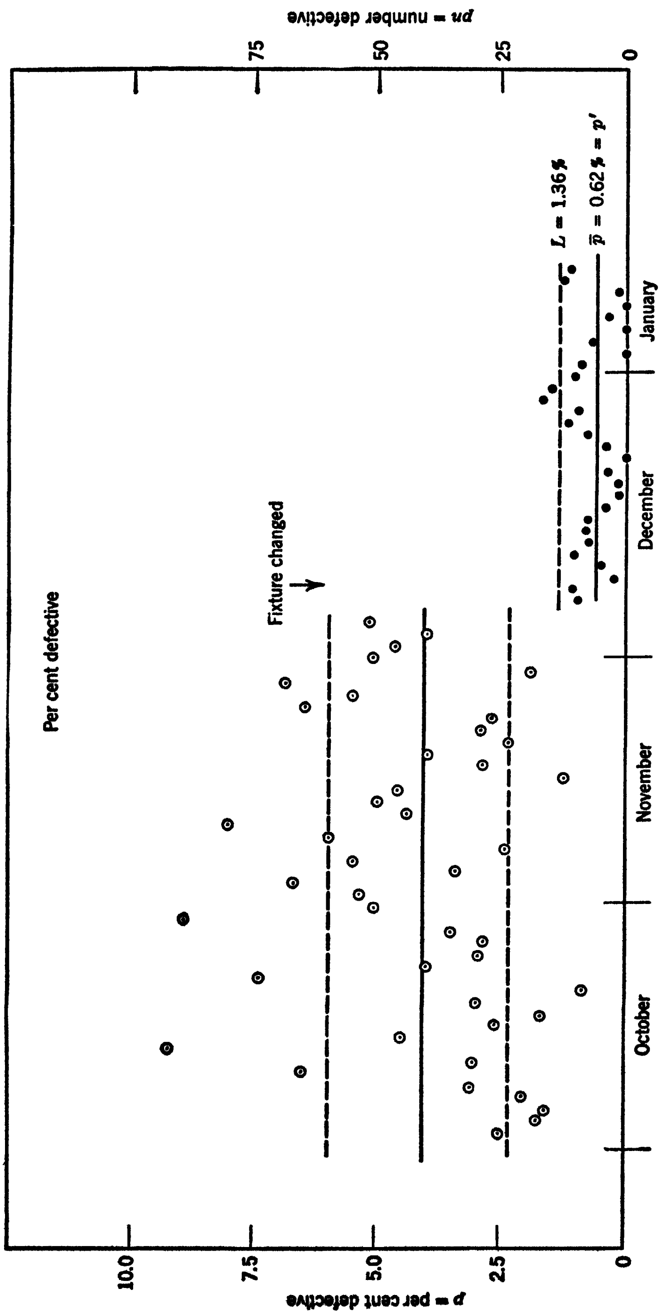


Figure 4. Tight lever

Table 10 and Figure 4 illustrate the method of calculating and plotting p charts with a constant sample size, using as an illustration the reversing mechanism of a ratchet wrench which is actuated by a lever on the outside of the upper cover. Sometimes the lever is assembled too tightly, so that it will not move at the touch of a finger, as it should.

From the daily production of a medium-sized ratchet, 1,000 completed assemblies were taken at random and tested for tight levers. The result appears on the left-hand side of Figure 4, and in Table 10. Variations were much larger than they should have been, as shown by many out-of-control points. Some unpredictable cause was apparently preventing a consistent quality of work from being done. A study of the assembly process soon revealed that the fixture used in welding the lever was poorly designed, so that consistent welds could not be obtained. Redesigning the fixture eliminated most of the trouble and reduced the average percentage of tight levers from 4.2 to 0.6 per cent.

In the second part of Figure 4 is shown a new p chart, calculated on the improved process. Two points in the latter part of December were out of control owing to a batch of faulty levers. Even then the 1.7 per cent and 1.5 per cent defective pieces found on those two days were fewer than were found most of the time on the old fixture.

pn CHART—CONSTANT SAMPLE SIZE

If the sample size is or can be made the same at each inspection, a pn chart may be more convenient to use than a p chart. On the pn chart the central line is $\bar{p}n$ = average fraction defective times number of pieces inspected. $\bar{p}n$ is therefore the expected number of defectives found in the samples. Control limits are calculated by the following formula:

$$\bar{p}n \pm 3\sqrt{n\bar{p}(1-\bar{p})}$$

These control limits determine the maximum and minimum number of defectives that might be found in any sample of size n as a result of chance sampling fluctuations alone, if the process is maintaining the average quality defined by $\bar{p}n$. p charts easily can be converted to pn charts by altering the scale and multiplying the central line and control limits by n. In Figure 4 the p scale is on the left-hand margin and the pn scale is on the right-hand margin. $\bar{p} = 4.19\%$ or 0.0419 becomes $\bar{p}n = 41.9$ defectives; the upper limit of 6.1 per cent becomes 61 defectives (at most) expected in a sample of 1,000; the lower limit of 2.3 per cent becomes 23 defectives (at least) expected in 1,000 assemblies.

TABLE 10

CALCULATION OF p CHART AND pn CHART—CONSTANT SAMPLE SIZE, TIGHT LEVER ORIGINAL PROCESS

(1) Date, 1942	(2) Sample Size, n	(3) No. Defectives, pn	(4) Per Cent Defective, p
Oct 3	1,000	25	2.5
4	1,000	18	1.8
5	1,000	16	1.6
6	1,000	20	2.0
7	1,000	33	3.3
10	1,000	65	6.5
11	1,000	30	3.0
12	1,000	92	9.2
13	1,000	45	4.5
14	1,000	26	2.6
17	1,000	17	1.7
18	1,000	30	3.0
19	1,000	8	0.8
20	1,000	74	7.4
21	1,000	41	4.1
24	1,000	29	2.9
25	1,000	28	2.8
26	1,000	35	3.5
27	1,000	90	9.0
28	1,000	51	5.1
Nov 2	1,000	53	5.3
3	1,000	67	6.7
4	1,000	34	3.4
5	1,000	55	5.5
6	1,000	24	2.4
9	1,000	60	6.0
10	1,000	81	8.1
11	1,000	44	4.4
12	1,000	50	5.0
13	1,000	46	4.6
16	1,000	12	1.2
17	1,000	28	2.8
18	1,000	40	4.0
19	1,000	23	2.3
20	1,000	29	2.9
23	1,000	27	2.7
24	1,000	65	6.5

TABLE 10 (*Continued*)

CALCULATION OF p CHART AND pn CHART—CONSTANT SAMPLE SIZE, TIGHT LEVER, ORIGINAL PROCESS

(1) Date, 1942	(2) Sample Size, n	(3) No. Defectives, pn	(4) Per Cent Defective, p
Nov 25	1,000	55	5.5
26	1,000	69	6.9
27	1,000	18	1.8
30	1,000	51	5.1
Dec 1	1,000	47	4.7
2	1,000	40	4.0
3	1,000	52	5.2
Totals	44,000	1,843	

Average per cent defective $= \bar{p}\% = 1{,}843/44{,}000 = 0.0419 = 4.19\%$

$$\bar{p}\% \pm 3\sqrt{\frac{\bar{p}\%(100 - \bar{p}\%)}{n}}\,\% = 4.19\% \pm 3\sqrt{\frac{4.19(100 - 4.19)}{1{,}000}}\,\%$$

$$= 4.19\% \pm 1.90\% = 6.1\%,\ 2.3\%$$

$$\bar{p}n = 0.0419 \times 1{,}000 = 41.9$$

$$\bar{p}n \pm 3\sqrt{n\bar{p}(1 - \bar{p})} = 41.9 \pm 3\sqrt{1000 \times 0.0419 \times 0.9581}$$

$$= 41.9 \pm 19.0 = 61,\ 23$$

that is, there should be in each sample of 1,000 levers not more than 61 defectives nor less than 23 defectives.

p CHART—VARYING SAMPLE SIZE

Sometimes it is not possible to maintain a constant sample size for inspection by attributes. Records of 100-per-cent inspection, of samples from lots of varying sizes, of receiving inspection, of process inspection often call for samples varying considerably in size. Because chance sampling fluctuations decrease as sample size increases, it is necessary to calculate new control limits for each point on the chart if the sample size is not constant. Limits are closer together for large samples than for small ones.

In Chapter 5, Case History and Chart XIII, will be found an illustration of a p chart with varying sample sizes. Table 12 shows the method of calculating the $\bar{p}$ line and the limits.

TABLE 11

CALCULATION OF p AND pn CHART—CONSTANT SAMPLE SIZE, TIGHT LEVERS, IMPROVED FIXTURE

Sample Size $n = 1{,}000$

(1) Date, 1942	(2) No. Defectives, pn	(3) Per Cent Defective, p
Dec 4	8	0.8
5	10	1.0
7	2	0.2
8	5	0.5
9	10	1.0
10	6	0.6
11	7	0.7
14	6	0.6
15	3	0.3
16	1	0.1
17	2	0.2
18	3	0.3
21	0	0.0
22	4	0.4
23	7	0.7
24	12	1.2
28	9	0.9
29	17	1.7
30	15	1.5
31	10	1.0
Jan 1943		
3	8	0.8
4	0	0.0
7	6	0.6
8	0	0.0
9	3	0.3
10	0	0.0
11	1	0.1
14	13	1.3
15	12	1.2
Total	180	

$N = \Sigma n = 29{,}000$

$\bar{p} = 180/29{,}000 = 0.0062 = 0.62\%$

$$\bar{p}\% \pm 3\sqrt{\frac{\bar{p}\%(100 - \bar{p}\%)}{n}} = 0.62\% \pm 3\sqrt{\frac{0.62(100 - 0.62)}{1{,}000}}\ \% = 0.62\% \pm 0.74\%$$

p limits $= 0.62\% \pm 0.74\% = 1.36\%, 0.0\%$

$\bar{pn} = 0.0062 \times 1{,}000 = 6.2$

$\bar{pn} \pm 3\sqrt{n\bar{p}(1 - \bar{p})} = 6.2 \pm 3\sqrt{1{,}000 \times 0.0062 \times 0.9938} = 6.2 \pm 7.4$

pn limits $= 6.2 \pm 7.4 = 13.6, 0$

TABLE 12

p CHART—VARIABLE SAMPLE SIZE

Hammer Operator Quality (Chart XIII)

(1) Date	(2) Sample Size, n	(3) No. Defective, c	(4) $c/n \times 100 = p(\%)$	(5) $3\sqrt{\bar{p}(1-\bar{p})}/\sqrt{n} = 35.25\%/\sqrt{n}$	(6) $\bar{p} \pm 3\sqrt{\bar{p}(1-\bar{p})/n}$ Limits	
Mar 1	34	0	0%	6.0%	7.4%	0.0%
2	47	0	0	5.1	6.5	0.0
5	48	0	0	5.1	6.5	0.0
6	68	1	1.5	4.3	5.7	0.0
7	55	0	0	4.7	6.1	0.0
8	42	0	0	5.4	6.8	0.0
12	127	0	0	3.1	4.5	0.0
13	118	0	0	3.2	4.6	0.0
14	20	2	10.0	7.9	9.3	0.0
15	52	0	0	4.9	6.3	0.0
16	27	0	0	6.8	8.2	0.0
16	14	0	0	9.4	10.8	0.0
19	12	0	0	10.1	11.5	0.0
19	52	0	0	4.9	6.3	0.0
20	54	0	0	4.8	6.2	0.0
21	35	0	0	6.0	7.4	0.0
21	17	0	0	8.5	9.9	0.0
21	31	1	3.0	6.3	7.7	0.0
22	67	3	4.5	4.3	5.7	0.0
23	72	2	2.8	4.2	5.6	0.0
24	53	4	7.6	4.8	6.2	0.0
24	28	2	7.1	6.7	8.1	0.0
27	133	3	2.3	3.1	4.5	0.0
28	85	2	2.4	3.8	5.2	0.0
30	65	3	4.6	4.4	5.8	0.0
Apr 2	36	0	0	5.9	7.3	0.0
3	121	1	0.8	3.2	4.6	0.0
4	19	1	5.3	8.1	9.5	0.0
4	9	0	0	11.7	13.1	0.0
5	54	0	0	4.8	6 2	0.0
6	56	0	0	4.7	6.1	0.0
10	53	1	1.9	4.8	6.2	0.0
11	53	0	0	4.8	6.2	0.0
12	45	0	0	5.3	6.7	0.0
13	45	0	0	5.3	6.7	0.0
14	38	1	2.6	5.7	7.1	0.0
16	25	0	0	7.0	8.4	0.0
17	37	0	0	5.8	7.2	0.0
18	13	0	0	10.2	11.6	0.0
19	62	4	6.4	4.5	5.9	0.0
22	24	0	0	7.2	8.6	0.0
25	48	0	0	5.1	6.5	0.0
26	48	0	0	5.1	6.5	0.0
27	54	0	0	4.8	6.2	0.0
28	6	0	0	14.4	15.8	0.0

Totals 2,202 31

$\bar{p} = 31/2{,}202 = 0.014 = 1.4\%$

$3\sqrt{\bar{p}(1-\bar{p})} = 3\sqrt{0.014 \times 0.986} = 0.3525 = 35.25\%.$

The limit for each point is based upon $\bar{p}$ for all the points and upon n for the individual points, but not upon the value of p for each point. Calculation of the limits therefore is accomplished most easily by separating the formula into two parts: $3\sqrt{\bar{p}(1-\bar{p})}$ which is a constant for all the points, and $\sqrt{n}$, which varies with each point. This method of calculating is used in Table 12.

THE CHART FOR p LIMITS

Calculation of p-chart limits may be tedious if done with pencil and paper, because of the necessity for extracting square roots. With a slide rule it is quicker. With the use of Figure 5, however, practically

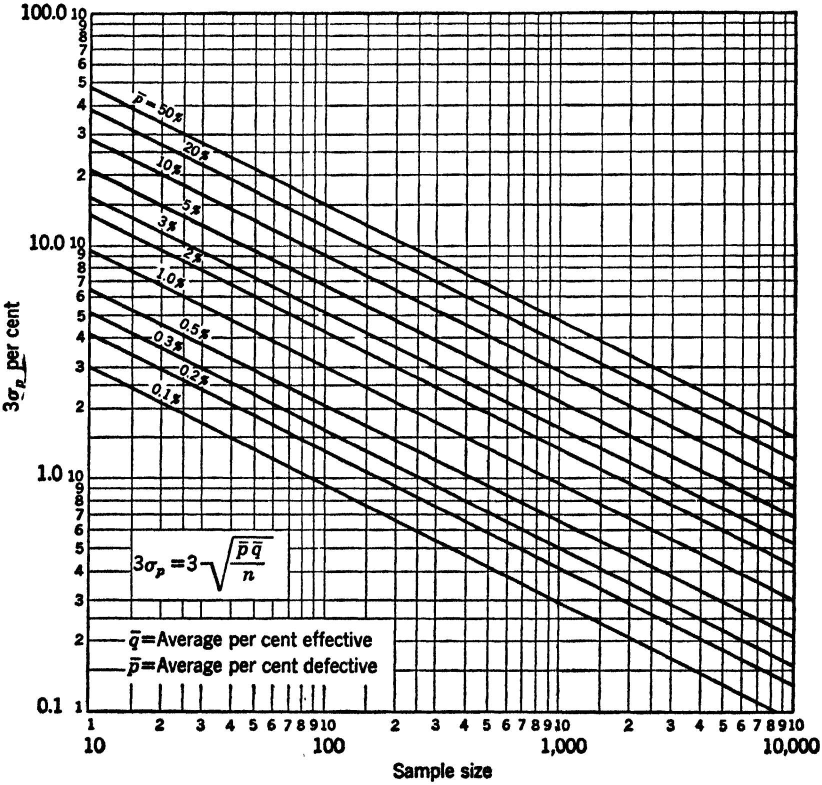


Figure 5. Chart for finding $3\sigma_p$ limits. Given $\bar{p}$ or $\bar{q}$ and n

all the mechanical difficulties of constructing a p chart are avoided. It enables an estimate of the limits to be made accurately enough for practical purposes.

Enter Figure 5 with the sample size, n. Go upward until the n ordinate intersects the diagonal $\bar{p}$ line at the value calculated from the data (interpolating if necessary). Go left from the intersection to the scale on the left-hand margin. Add and subtract this value from the calculated $\bar{p}$ in order to get the control limits. If the lower limit is negative, take zero as the lower limit.

Illustration: Suppose that at least 20 p points have been derived from inspection data, and that the $\bar{p}$ is 5.0%, with a sample size n of 100. Start with the 100 ordinate on the scale at the bottom of the chart, go upward to the diagonal $\bar{p} = 5.0\%$ line, and then over to the left-hand scale. The value is approximately 6.7 per cent. Add and subtract this from $\bar{p}$. The limits are 5.0 per cent $\pm$ 6.7 per cent, or 11.7 per cent and 0.0 per cent. Actual calculation shows:

$$\bar{p}\% \pm 3\sqrt{\frac{\bar{p}\%(100 - \bar{p}\%)}{n}} = 5.0\% \pm 3\sqrt{\frac{5.0 \times 95.0}{100}}\%$$
$$= 5.0\% \pm 6.5\% = 11.5\%, 0.0\%$$

The discrepancy in estimate of 0.2 per cent between the chart and the calculation is not large enough to make any practical difference when such a p chart is put to actual use.

In the preceding illustration, if $\bar{p}$ were 4.0 per cent and n were 155, it would be necessary to interpolate. Go up from the $n = 155$ ordinate to approximately halfway between the $\bar{p} = 3.0\%$ and $\bar{p} = 5.0\%$ diagonal lines, then over to the left-hand scale. The result is, 4.0% $\pm$ 4.6% = 8.6%, 0.0%. Calculation gives values of 8.7 per cent, 0.0 per cent.

THE MEANING OF A p CHART

An out-of-control point on a p or pn chart indicates the presence of an assignable cause: the lot is probably different in quality from the others. In essence, the control limits give the maximum and minimum per cent of defectives that may be found by chance in a sample of size n from a lot of presumably $\bar{p}$ quality. If too many or too few defects are found, any one or all of three broad types of assignable causes may exist:

(*a*) The lot quality may be different; that is, something may have happened to the manufacturing process.

(*b*) The inspection procedure may have changed; carelessness, varying standards, errors in gauges, or nonrandom selection of the sample may be the cause.

(*c*) The product inspected may be physically not the same as that inspected previously; soft bolts might be mixed with hard ones, or

two types of raw material might be present, or some of the articles may have been made by one process and some by another process.

As to which of these three types of causes may be to blame, the control chart can only point out that *some change probably has occurred.* What has occurred and why can be discovered only by an investigation, using the clues provided by statistical analysis.

REAL AND APPARENT DIFFERENCES IN QUALITY

How a control chart can distinguish between real and apparent differences in quality is illustrated by the following analysis of a welding problem.

In a certain factory there was a battery of 29 welding machines, all doing the same general type of butt welding. A good many defective welds were showing up. The question arose as to what was the source of the trouble.

A sample of 50 welds was selected at random from a day's production off each machine and was tested to determine shear strength. Table 13 shows the results of the test.

It was worth while to know that the process averaged 10.2 per cent defective, but that alone did not help much in detecting where the bad work was being done. Machine 10, with $^{10}/_{50} = 20\%$ defective might be considered significantly worse than the average; machine 8 with $^{1}/_{50} = 2\%$ defective might be considered better than average. These apparently obvious conclusions should be tested statistically before being accepted as facts. If 10 weak welds in 50 is too high an amount of bad work, how about 9 or 8 or 7: are they also significantly bad?

This problem can be viewed from another angle. If each machine is judged by the single sample of 50 pieces, what is the probable per cent defective from each machine? This question can be answered in part by using the formula, $\sigma_p = \sqrt{\frac{p(1-p)}{n}}$ where σ_p is read, "the standard deviation of a percentage"; p is the apparent fraction defective shown by the sample; $(1 - p)$ is the apparent fraction good; and n is the number of observations in the sample. If we use $3\sigma_p$, as in calculating control-chart limits, and substitute the numerical values for machine 8: $n = 50$, $p = \frac{1}{50} = 0.02$, $1 - p = 0.98$,

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.02 \times 0.98}{50}} = \sqrt{0.000392} = 0.02$$

and

$$3\sigma_p = 3 \times 0.02 = 0.06$$

TABLE 13

PERCENTAGE OF DEFECTIVE WELDS

Machine No.	No. Welds Tested	No. Defective Welds Found
1	50	3
2	50	6
3	50	2
4	50	8
5	50	7
6	50	5
7	50	4
8	50	1
9	50	8
10	50	10
11	50	4
12	50	4
13	50	4
14	50	7
15	50	9
16	50	7
17	50	2
18	50	5
19	50	6
20	50	3
21	50	5
22	50	5
23	50	5
24	50	3
25	50	5
26	50	2
27	50	4
28	50	9
29	50	5
Totals	1,450	148

Average per cent defective = 148/1,450 = 10.2%

The actual fraction defective of machine 8, as judged from the sample of 50 welds, almost certainly is no worse than $p + 3\sigma_p = 0.02 + 0.06 = 0.08$. That is, the lot actually may be as much as 8 per cent defective.

Similarly, for machine 10,

$$n = 50, \quad p = {}^{10}\!/_{50} = 0.20, \quad 1 - p = 0.80$$

and

$$3\sigma_p = 3\sqrt{\frac{p(1-p)}{n}} = 3\sqrt{\frac{0.20 \times 0.80}{50}} = 0.17$$

Lot 10 therefore could be as bad as $0.20 + 0.17 = 0.37 = 37\%$.

Neither of these answers regarding lots 8 and 10 is very satisfactory because of the wide variation in possible quality. The estimates are so wide in range that they do not serve a very useful purpose. The unreliability is due to the small sample of only 50 pieces. Closer estimates could be made by increasing the sample size.

The formula for the standard deviation of a percentage σ_p is nevertheless useful in answering the question as it was first posed: "Which of the 29 machines is significantly better or worse than the average? Which of them is primarily responsible for the more-than-permissible amount of bad work?"

Let the average of all 29 machines' quality be $\bar{p}$. Then,

$$\bar{p} = 0.102, \quad 1 - \bar{p} = 0.898, \quad n = 50$$

$$\sigma_p = \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} = \sqrt{\frac{0.102 \times 0.898}{50}} = 0.043$$

$$3\sigma_p = 3 \times 0.043 = 0.129$$

and

$$\bar{p} \pm 3\sigma_p = 0.102 \pm 0.129 = 0.231 \text{ and } 0$$

That is, given a process 10.2 per cent bad, if samples of 50 are taken from various lots, the samples may be expected to vary between 0 and 23.1 per cent defective. Samples of 50, therefore, might have the following minimum and maximum number of defects:

$$0\% \text{ of } 50 = 0 \text{ pieces defective}$$

$$23.1\% \text{ of } 50 = 11.55 \text{ pieces defective}$$

Even the worst machine (No. 10) actually had only 10 defectives in its sample, whereas 11 defectives in the sample could have occurred by chance *if the machines actually had been producing* 10.2 *per cent bad work.* Clearly, on the basis of the evidence available, none of the machines justly can be accused of exceptionally bad work.

This problem will be recognized as an application of p- and pn-chart techniques, with the data analyzed by machines instead of in order of time. If the $3\sigma_p$ limits are considered too far apart and closer limits such as $2\sigma_p$ are used, the maximum variation due to chance comes out at $\bar{p} + 2\sigma_p = 0.102 + 2 \times 0.043 = 0.188$, or 9.4 defective pieces in a sample of 50. Since machine 10, with 10 defectives, is barely outside the $2\sigma_p$ limit, it would be rash to conclude that anything is significantly wrong with no. 10. Such a sample would be found about one time in 20 from a 10.2 per cent bad process. Since there were

29 samples taken, one of them would be expected to show as many as 10 defectives. No. 10 may have just happened to be that one.

The conclusion drawn from this study was that the source of bad work probably lay not in any individual machine but in some flaw in the process itself—in raw materials, design, or workmanship.

SAMPLE SIZE

In a p or pn chart the average line, $\bar{p}$ or $\bar{p}n$, in the inspected samples should be as close as possible to the true but unknown average of the uninspected part of the product. If the samples on which the limits are calculated do not reflect the true condition of the product, the chart may become dangerously invalid. A frequent error in using p and pn charts is the selection of a sample size either too large or too small. If the sample is too large, an accuracy greater than necessary will be achieved at an excessive cost. If the sample size is too small, inadequate or unreliable information may result.

In Table 10, for instance, a sample of 1,000 ratchets was inspected each day. As is shown later, it was unnecessary to have as large a sample as that in order to get sufficiently reliable estimates of the process average. If 135 ratchets instead of 1,000 had been inspected, the $\bar{p}$ on the chart would have been very little different from the 4.19 per cent actually found. Control limits would have been wider, but nevertheless sufficiently sensitive to serve the purpose of the chart. Seven times as much inspection was done as was necessary to discover and eliminate the cause of the trouble.

In Table 12 and Chart XIII, on the other hand, the samples were too small to be satisfactory. Of 45 inspections made, 30 showed no defects. Each of these zeros counted as a perfect sample in calculating the $\bar{p}$ of 1.4 per cent, but probably none of the work from which the 30 samples came was really perfect. Some bad work certainly existed which the samples were too small to detect. The $\bar{p}$ of 1.4 per cent therefore was probably *underestimated.* It may have been higher than 1.4 per cent, but how much higher it was impossible to know.[1] A good deal of valuable information was lost; however, by putting the inspection records in the form of a control chart, the maximum benefit was gained from the data that were available.

A sample too large is generally uneconomical; a sample too small is

[1] In the process described by Case History XIII it was uneconomical, in fact practically out of the question, to take adequate samples. The smaller samples that were taken, though statistically unsatisfactory, proved nevertheless to be extremely valuable.

dangerous unless its limitations are realized and allowed for. A happy solution and a practical answer to this problem can be suggested from the author's experience:

The sample should be large enough so that at least nine times out of ten one or more defectives will be found.

Suppose that there is reason to believe, from previous experience, that a product is coming from the factory with an average of 5 per cent defective. How large a sample size n would be required to find at least one bad piece in the sample nine times out of ten?

In order to answer that question, visualize a pn chart with a central line at $\bar{p}n$ and a lower-limit line at $\bar{p}n - 2\sigma_{pn} = 1.$[2]

Solving the equation $\bar{p}n - 2\sqrt{n\bar{p}(1 - \bar{p})} = 1$ for n gives

$$\sqrt{n} = \frac{\sqrt{2 - \bar{p}} + \sqrt{1 - \bar{p}}}{\sqrt{\bar{p}}}$$

This equation can be plotted as an approximately straight line on log-log graph paper, as in Figure 6. Enter the chart with the value of $\bar{p}$ on the left-hand scale, go across to the diagonal line, and read down to the value of n on the bottom scale. This is the minimum sample size for a good p or pn chart and should be used wherever it is economical and practical to do so.

A similar question arises in connection with the upper limit of a p or pn chart. Suppose that, from previous experience, a product is known to be 5 per cent defective on the average. Suppose that a quality limit not worse than 10 per cent is desirable, and that a 10 per cent bad product will not be permitted more than two times in a thousand. This is equivalent to setting a value of $\bar{p} + 3\sigma p = 10\%$. Solving the equation $\bar{p} + 3\sqrt{\dfrac{\bar{p}(1 - \bar{p})}{n}} = L$ gives

$$n = \frac{9\bar{p}(1 - \bar{p})}{(L - \bar{p})^2}, \quad \text{where } L \text{ is the quality limit.}$$

This value can be read from Figure 5 by looking up $(L - \bar{p}) = 3\sigma p$ on the left-hand margin, going across to the diagonal $\bar{p}$ line (interpolating if necessary) and then down to the value of n on the bottom scale. For $\bar{p} = 5.0\%$ and $L = 10.0\%$, $L - \bar{p} = 5.0\%$. Start at 5.0 per cent on the left-hand scale, go across to $\bar{p} = 5.0\%$ diagonal line, and down to $n = 175$, approximately.

[2] A $2\sigma_{pn}$ rather than a 3 σ_{pn} limit is used because the probability of failing to get a defective is to be about 0.10 instead of in the region of 0.005.

There are therefore two sample sizes from which to choose. In the present case, when $\bar{p} = 5.0\%$ and $L = 10.0\%$, n may be 112 or 175. *Whichever n is larger should be used for the control chart.*

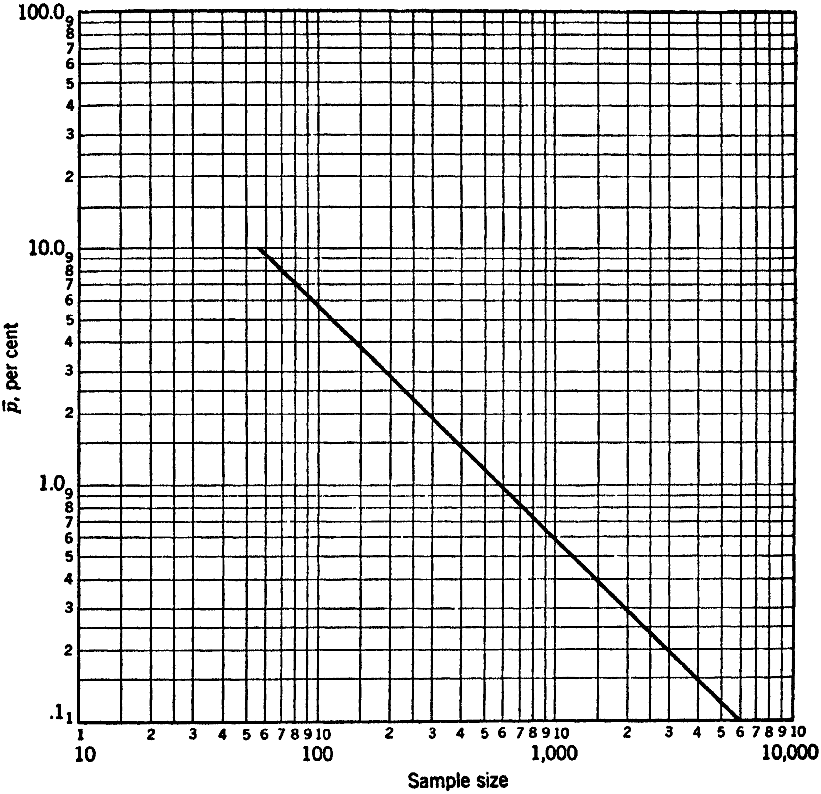


Figure 6. Chart for determining minimum sample size.
$[\bar{p}n - 2\sqrt{n\bar{p}(1 - \bar{p})} = 1]$

SETTING STANDARD VALUES

A number of schemes are available for controlling the outgoing quality of factory products by sampling inspection. These plans envision inspection as a sieve through which the *finished product* must pass in order to prevent shipment of unacceptable goods to the consumer. The Dodge–Romig tables,[3] the Army Ordnance tables,[4] Simon's I_Q

[3] H. F. Dodge and H. G. Romig, *Sampling Inspection Tables*, John Wiley & Sons, Inc., New York, 1944.

[4] The Army Ordnance tables provide for single- and double-sampling plans similar to the Dodge–Romig tables. They have not been published.

charts,[5] Working's Poisson chart [6] are some of the better-known sampling plans.

These plans perform a useful function, but they do not *prevent* the damage before it is done. Not until the *process* is controlled at the desired outgoing quality level can the manufacturer realize the maximum benefits of good production and engineering practice. Control of the process is basic to control of quality. It is therefore essential to set process standards and to be sure that they are met consistently. For this purpose p and pn charts serve well.

First, of course, comes a study and analysis of the process,[7] including action designed to improve and stabilize manufacturing methods. Next, a standard average per cent defective p' should be determined, such that p' the standard is equal to or larger than $\bar{p}$ the actual process average. Third, the desired outgoing quality limit L should be decided upon; it must be larger than p' and should be as liberal as possible for the sake of economy.

The minimum sample size necessary for maintaining the process within specified p' and L limits can be read off Figure 5 as previously explained, provided such a sample is large enough to show at least one defective (use Figure 6).

Keeping control in this manner requires three things not demanded by other sampling plans:

(*a*) A knowledge of what the process is doing and can accomplish; that is, a known value of $\bar{p}$.

(*b*) A realistic standard or acceptable average quality p' based upon and equal to or larger than the known $\bar{p}$.

(*c*) A realistic quality limit L larger than p' and if possible double p', which is in effect an upper control-chart limit:

$$L = p' + 3\sigma_p = p' + 3\sqrt{\frac{p'(1 - p')}{n}}$$

In return for these requirements the method provides:

(*a*) Strong pressure upon engineers and production men to improve the process as rapidly and extensively as necessary for setting practical and desirable p' and L values.

(*b*) A proved tool for maintaining quality once statistical control has been achieved, namely, the control chart.

[5] L. E. Simon, *An Engineer's Manual of Statistical Methods*, John Wiley & Sons, Inc., New York, 1941.

[6] Holbrook Working, *A Guide to the Utilization of the Binomial and Poisson Distributions*, Stanford University, 1943.

[7] See Case Histories I–V, Chapter 5.

(*c*) An economical and adequate inspection procedure.

(*d*) Reduced inspection costs.

Reduced inspection of a controlled process is one of the major benefits derived from the use of control charts. A suggested procedure for reduced inspection is as follows:

1. When 20 successive points have remained in control, reduce the sample size to one-fifth the previous size, and recalculate the control limits for the new smaller n. Each point on the chart thereafter must be in control inside the new limits; and five successive inspections grouped together must be in control within the old limits.

2. Any point out of control on the old limits requires detail or sorting inspection of the lot from which the sample came and an investigation of the process for assignable causes.

3. Any point out of control on the new limits (with small n) requires four more small samples to be taken immediately out of the suspected lot. If all five samples combined meet the old test on the old limits, the lot may be accepted and reduced inspection may continue. If the combined five small samples fail to meet the old test, detail inspection of the lot is required, with a return to the old large sample size until another 20 points have remained in control.

This procedure, by putting a high premium on a controlled process, compels prompt investigation and correction whenever the quality of the *process* begins to slip. If the process is kept in control, practically no detail inspection will be required, and sampling inspection can be performed with economy as well as with a high degree of quality assurance.

AN ILLUSTRATION

The chart on tight levers, Figure 4 and Tables 10 and 11, illustrates the control-chart inspection procedure already outlined. In its original state the process had an unsatisfactory average quality of $\bar{p} = 4.19\%$. Changing the fixture reduced this to 0.62 per cent, an acceptable $\bar{p}$. The standard quality therefore was set at $p' = 0.6\%$.

As to sample size, reference to Figure 6 reveals that a sample of only 135 pieces was required for establishing the first $\bar{p}$ of 4.19 per cent with a practical degree of accuracy. The sample size actually was 1,000, more than seven times as large as was needed. Use of Figure 6 in this case would have greatly reduced the inspection cost.

With the improved process $\bar{p}$ of 0.62 per cent, reference to Figure 6 shows that a sample of about 930 pieces was required. The sample size of 1,000 was therefore not much too large. Management decided to set a standard average quality of 0.6 per cent, with an upper limit

of 1.5 per cent. Entering Figure 5 on the left-hand margin with $L - p' = 1.5\% - 0.6\% = 0.9\%$, going across to the diagonal $\bar{p} = 0.6\%$ (interpolated between 0.5% and 1.0%), and reading downward gives a sample size required for holding the standards of about 600 pieces. Of the two required sample sizes, the one from Figure 6 of 930 pieces was the larger. Management decided in this case to use $n = 1{,}000$ for convenience in computing, since it was only slightly larger than the required sample. The upper control limit therefore was set as shown on the control chart (Figure 4, Table 11), giving a maximum quality limit of 1.36 per cent which was slightly better than the 1.5 per cent standard.

After the improved process was started, 17 successive points remained inside the control limits, but the 18th and 19th points went out of control. Samples of 1,000 pieces continued, therefore, to be used, because the process had not yet qualified for reduced inspection. The next 20 points (only 10 of them shown on Figure 4), were in control. Beginning with the 40th lot, samples of only 200 pieces were taken; the reduced samples were grouped by fives (nos. 40–44, 45–49, 50–54, etc.), so that each group contained 1,000 pieces; the total rejects found in each group was calculated as a percentage of 1,000 and was plotted on a continuation of Figure 4. Each of the grouped points had to be in control within the upper limit of 1.36 per cent; if it fell outside that limit, it was detailed, and inspection returned to the large $n = 1{,}000$ sampling procedure until another 20 points were in control.

In the meantime, while reduced inspection was in force, new control limits were calculated for the sample size of 200,

$$p' \pm 3\sigma_p = 0.6\% \pm 3\sqrt{\frac{0.6 \times 99.4\%}{200}} = 0.6\% \pm 1.6\% \doteq 2.2\%, 0\%$$

and each small sample was required to be in control on these limits; if a small-sample point went out of control, four more small samples were taken from the lot; if the combined sample of $5 \times 200 = 1{,}000$ showed a percentage defective of more than 1.36 per cent, inspection reverted to the large-sample procedure until another 20 points had remained in control.

SOME PRECAUTIONS IN TAKING SAMPLES

In using p and pn charts, as well as $\bar{X}$ and R charts, it is essential that care be taken in choosing the samples from rational subgroups chosen upon the best available apriori information about the assign-

able causes that are likely to be present. It must be kept in mind that the purpose of the control chart is to detect the presence of assignable causes and that the proper method of sampling depends on what you are trying to discover. How best to sample in a given case is a question that should be left to one thoroughly familiar with the theory of sampling in control-chart work, a subject beyond the scope of this book. In other words, no fool-proof rules can be laid down that will apply to all cases. However, the author has found from experience that a person who observes the following precautions will usually succeed in doing a reasonably satisfactory job. At least a consideration of these precautions will give the reader an indication of some of the factors that should be considered.

1. If the product comes in trays, tubs, or bins, be sure that at least a few articles come from every part of each container. Do not choose all the articles in the sample from one spot or from one container.

2. Dig down into the container. Take some pieces from the top, some from the middle, and some from the bottom of each container.

3. If the product comes on a belt, take samples at convenient or economic intervals during the shift or run. Do not take the whole sample from a single short period of time.

4. Study the process to determine whether or not randomizing operations occur. Such study may uncover profitable times and methods of inspection that will considerably reduce inspection costs.

5. Be sure that the inspection operation—whether visual, gauging, or testing—always is performed in the same way on every piece in the sample.

6. For visual inspection have definite objective standards of finish, appearance, shape, and so on.

7. Check or calibrate all inspection instruments before beginning the inspection. A small fault in the instrument will introduce a constant error that may invalidate the records.

8. Do not mix different products. Each type of electric bulb, for instance, should be inspected separately. Similar products sometimes may be grouped if they are produced by the same process and if the resulting chart looks reasonable. Here common sense and statistics should be combined in order to assure a useful workable chart.

9. The person who selects the sample should understand thoroughly the need for randomness and the physical methods of getting it. To this important task only those specially trained in sample taking should be assigned. Otherwise the inspector, especially a good inspector, unconsciously will tend to select bad pieces; quite naturally, since he earns his living doing just that. A good inspector may not be able

to take a truly random sample unless he has been carefully trained to do it.

10. A chart out of control does not necessarily point to an unreliable *process*. The assignable cause also may be found in:

(*a*) Nonrandom samples.
(*b*) Uncertain inspection standards.
(*c*) Poorly designed or adjusted inspection instruments.
(*d*) Poorly supervised, poorly trained, or careless inspectors.
(*e*) Different methods of inspection.
(*f*) Different products mixed together.
(*g*) Incorrect recording of data and errors in calculation.

This is not a dissertation on inspection techniques; the preceding suggestions merely help to translate the essential idea of randomness into terms of physical operations. The theory upon which the p and pn charts are built assumes that each defect is *independent* of each other defect. In manufacturing operations, on the other hand, one defect is very apt to be followed by others due to the same cause, as when tooling or setup is poor and a succession of defective articles is produced. Only by taking care that the sample is selected at random can the physical world of production cause and effect, where defects are not independent, be made to fit the assumptions implicit in the control-chart technique.

By taking care to bring facts and theory into agreement, the manufacturer is able to take advantage of the benefits offered by the use of control charts: basic improvement in his operations, greater efficiency, more assurance of quality. These things have a definite and often large monetary value. It is therefore well worth while for the factory executive to design his inspection procedure scientifically and to provide engineering "detectives" for running down the clues discovered by his control charts. In the next chapter are given some illustrations of how this procedure works.

CHAPTER 5

CASE HISTORIES OF CONTROL CHARTS

Each of the case histories in this chapter has been selected to illustrate a particular application or a specific aspect of the control-chart technique. The author is indebted to many individuals and firms for the material presented here. Because they sometimes contain trade secrets, most of the charts and accompanying stories have been suitably camouflaged in order to avoid the release of confidential information. In no case, however, has the point at issue been camouflaged. It is hoped that these case histories will be helpful to those who wish to make practical use of control charts in their own plants, both by clarifying the techniques and by stimulating the imagination.

Charts I–V illustrate some applications of the p chart to process inspection, final inspection, and receiving inspection. Charts VI–X demonstrate a few of the situations to which $\bar{X}$ and R charts have been profitably applied. Charts XI–XV cover studies of an indirect nature, such as personnel problems, cost accounting, and production scheduling. These case histories are not intended to be an all-inclusive demonstration of control-chart practice, since they are limited to the author's personal contacts. They merely illuminate a few of the situations where control charts have proved in actual use their economic value, their power to earn money for a manufacturing business.

CASE HISTORY I. p CHART—PROJECTING COVER

The specifications for a metal container required the cover to be perfectly flush with the edge of the box. In assembly, however, projections of the cover above the edge or vice versa ran up to 0.00125 inch. It was necessary to perform an extra polishing operation in order to meet the "flush" specification. When the specification was changed to permit a 0.001-inch projection, most of the extra polishing work was eliminated.

Chart I illustrates the effect upon rejections of a realistic specification, one adapted to the process. The first part of the chart shows the result of the old process and the "flush" specification. After the

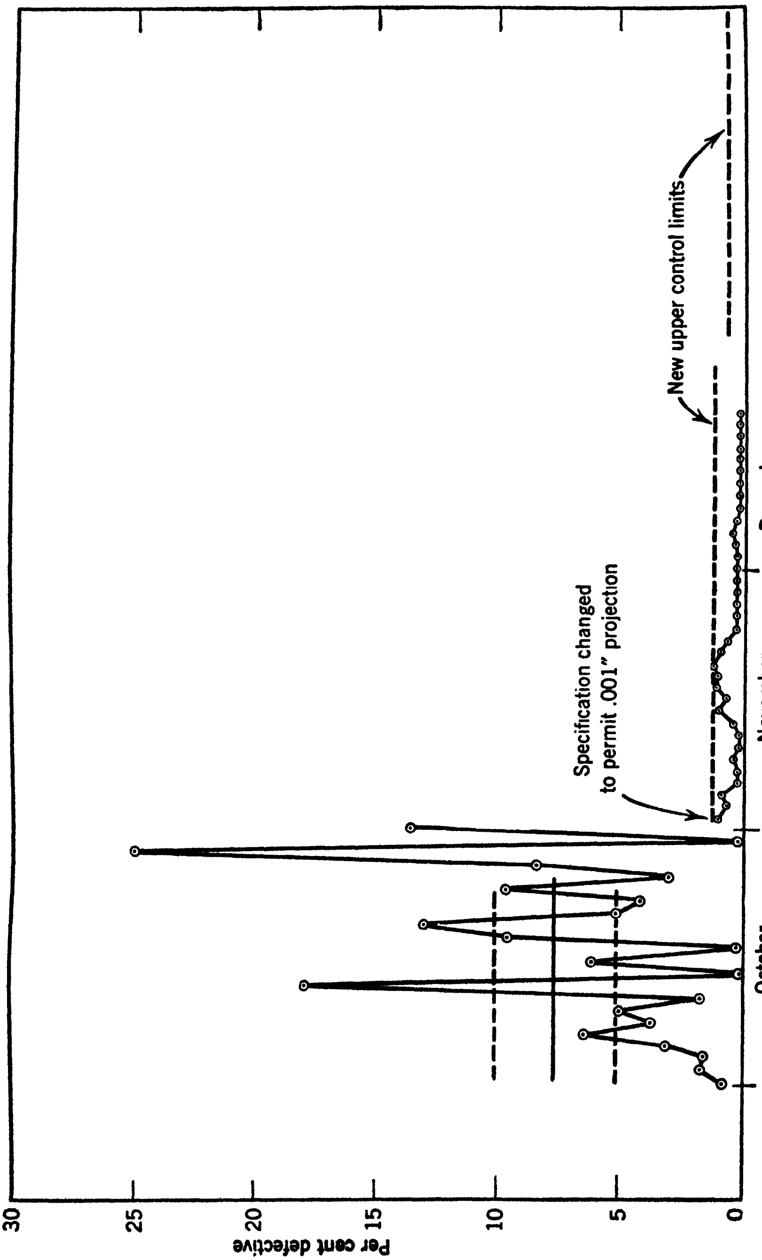


Chart I. Control chart for per cent defective. Projecting cover

specification had been liberalized to fit in with the process, results were as shown on the second part of the chart.

The data for this chart were derived from daily records of the number of containers that had to be preassembled and specially ground, as a percentage of the total number assembled. That is,

$$\text{Each daily point} = \frac{\text{no. specially ground}}{\text{no. assembled}} = p\%$$

Note: In the chart, three days show 0 per cent specially ground. On these three days no covers were assembled because experiments were in progress on the new specifications.

CASE HISTORY II. *pn* CHART—CONTROL OF INSPECTION

Scientific men long have realized that, no matter how accurate their instruments may be, there is always present in any series of measurements a residual human error of observation. Machines can be made practically perfect, but men never. The human-error element cannot be eliminated. It *can* be minimized and controlled. Chart II shows how the personal error in an inspection operation was held to small and predictable limits by using a control chart.

As a result of machining operations on a certain type of wrench, burrs frequently appeared on the handles. These burrs had to be ground off before the wrenches were acceptable to mechanics. The wrenches were therefore 100 per cent inspected by being passed over a belt one by one; a girl inspector examined them and threw out those with burrs.

Complaints occasionally were made that wrenches with burrs on them were being shipped to customers. In order to determine what basis there was for such complaints, samples were collected from the 100-per-cent-inspected wrenches throughout the day so that at the end of the day approximately 400 had been sample-inspected. A record was kept of the percentage found each day with burrs on the handle. Chart III shows the result graphically.

Apparently, before the chart was started, the 100 per cent inspection had been letting as high as 5 per cent burrs slip by into the warehouse. During the first week of the chart, the inspection improved rapidly, showing, on the seventh day, no burrs in the sample of 400. Thereafter the inspection, as tested by the sample check, proved to be consistently good, though of course not perfect. Calculations for days 7 through 26 showed a $\bar{p}n$ of 2 defectives with limits of 6 and 0.

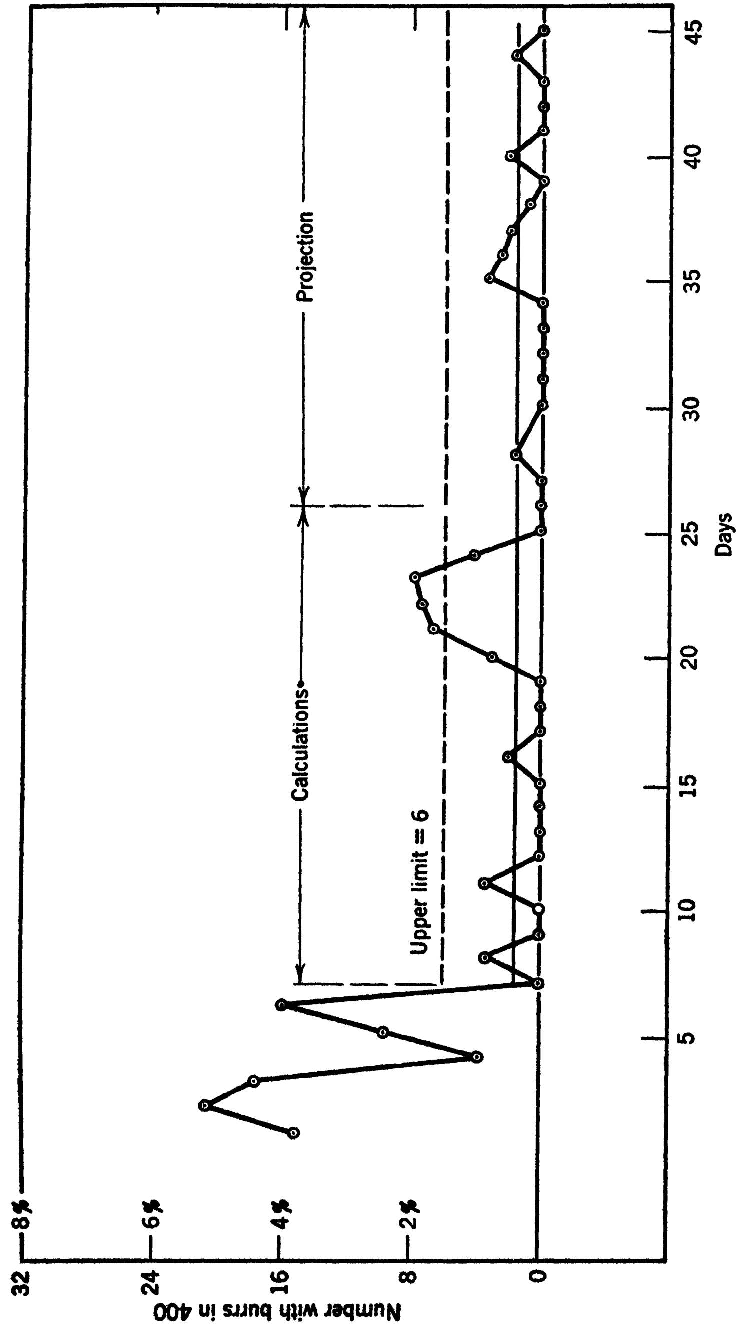


Chart II. Sample check on 100 per cent inspection. Burrs on handle. Sample of 400 taken after 100 per cent inspection

In other words, the 100 per cent inspection missed perfection on the average by 0.5 per cent, and almost never missed more than 1.5 per cent.

Under the circumstances this was considered a reasonable amount of human error. The upper control limit for samples of 400 taken *after* 100 per cent inspection was set at 6, with an average of 2.

Three points—the 21st, 22d and 23d days—were out of control on the new limit. It happened that on those particular days a new inspector was being broken in. It took three or four days to train her properly; thereafter her work maintained a consistently satisfactory level.

CASE HISTORY III. p CHART—CONTROL OF INSPECTION BY SHIFTS

Probably the most extreme case of inspection control by using control charts came to the author's attention in a plant making radar equipment. An electrical test was made on every tube produced in order to assure its performance in the field. Although failure in use normally would not endanger life or property, if the tube did not function at a critical moment, valuable time would be lost in repairing the equipment. A defective tube, therefore, although not critical, was a major fault because of the time and expense required to repair it.

When numerous failures were reported, the manufacturer's chief inspector suggested that an analysis of inspection records be made for the preceding four months. This was done and showed an average of 4.1 per cent rejects at the test, with a high of 17.7 per cent and a low of 0.0 per cent. Such a wide variation seemed inexplicable, and so control limits were calculated, using for n the average number tested per shift—1,250 tubes. Control limits came out at 5.8 per cent and 2.4 per cent. The astonishing thing was that almost none of the points were inside the limits: only one lot out of 27 tested fell inside; all the other 26 individual values of p were outside the control limits, either too high or too low. A review of the records showed immediately that all the low points had been tested by the first-shift inspectors, all the high points by the second shift. This situation is visualized on Chart IIIa.

A prompt investigation revealed that the first-shift inspector foreman had not instructed his men properly as to how to read the measuring instruments. The test was quite complicated, and correct readings were essential. The foreman, asked to make a test, proved that he himself did not know how to carry it out. His excuse was that when

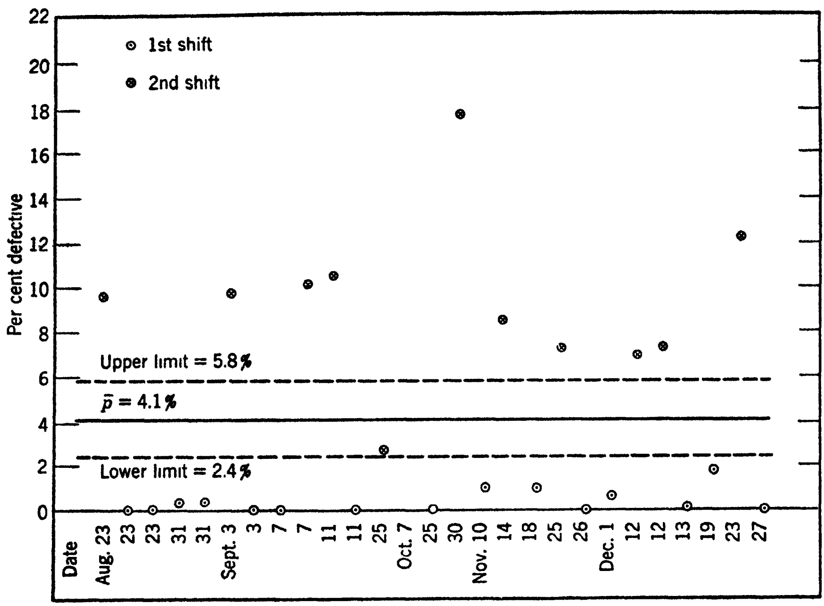


Chart IIIa. Control of inspection by shifts. Radar tube

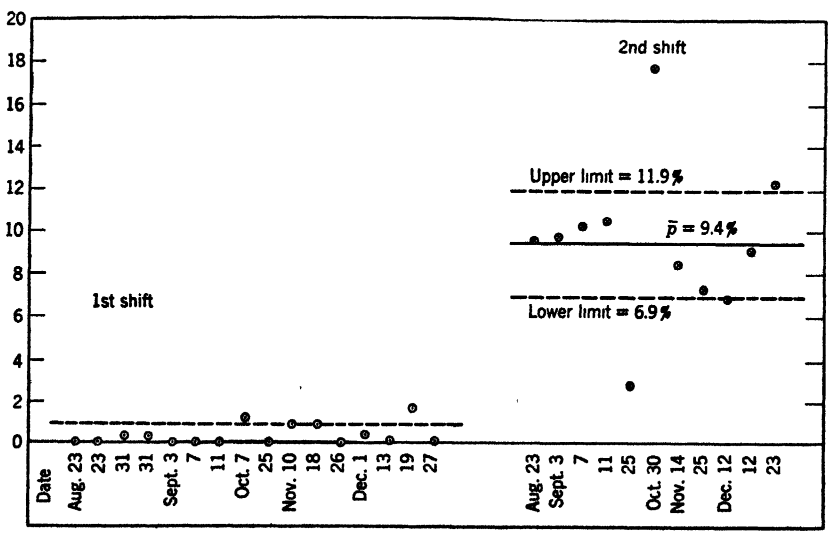


Chart IIIb. Inspection of radar tubes by shifts

he had first started as a test inspector the testing machine was a different one and that he had never been told how to operate the new machine.

In the meantime two separate charts had been made up, one for each shift. Lack of control in each case indicated lack of control in the process, in addition to the inspection fault. A further examination of the second-shift inspection records provided clues, in the shape of reasons for rejections, that led ultimately to a marked drop in the percentage of test failures.

Chart IIIb shows the segregation of the data by shifts and the two control charts derived from the breakdown.

Author's comment: Any point on a p chart or pn chart that falls below the lower control limit, indicating a quality better than expected, should be viewed with suspicion. Often such a "good" out-of-control point does not mean good work, but does mean poor inspection.

CASE HISTORY IV. p CHART—CHROME PLATING

Chart IV refers to a chrome-plating problem. The plating was done by a subcontractor. Chart IV records graphically the results of the prime contractor's daily receiving inspection. Between 2,000 and 2,500 plated parts were inspected each day, with an average of 2,280. Each point on the chart represents one day's inspection.

The central line and limits, $\bar{p} \pm 3\sqrt{\frac{\bar{p}\bar{q}}{n}}$,* as shown on the chart, were calculated after excluding the three excessively high points. Expressed as percentages, $\bar{p}$ was 2.1 per cent, $\sqrt{\frac{\bar{p}\bar{q}}{n}}$ was $\sqrt{\frac{2.1 \times 97.9}{2{,}280}}$ per cent or 0.32 per cent, and the limits worked out at $2.1\% \pm 3 \times 0.32\% = 3.1\%$ and 1.1%.

Before these limits were calculated, the three excessively bad days had been investigated. When the subcontractor saw this chart he was very much interested. Questioning his foreman and looking back over his production records, he discovered that one or two days before each high point a new rack had been used, but no change in the time of immersion in the plating solution had been made. A few experiments then were undertaken, which proved conclusively that a new rack, offering less resistance to the current, deposited a heavier plate during its first day than on succeeding days. The foreman, who had been

* $\bar{q} = 1 - \bar{p}$.

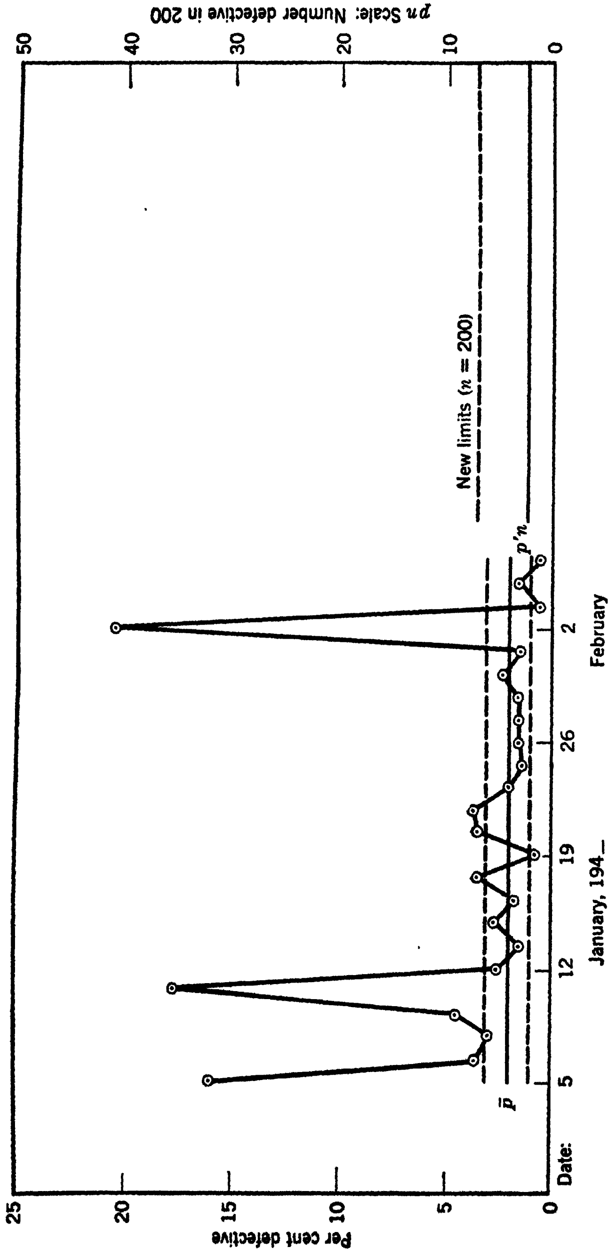


Chart IV. Control chart for per cent defective (oversize). Chrome plating

employed only a short time, had been unaware of the effect of a new rack. When he was told about it, he took special pains with his setup on new rack days and had no more trouble on that score.

Upon receiving the subcontractor's report on this investigation, the prime contractor extended the limits on the chart for a month in advance. At the end of the month a new $\bar{p}$ and new limits were calculated, and a *sample* receiving inspection of 200 pieces per day was set up. The new $\bar{p}$ became the standard p' (data not shown on the chart) at 1.2 per cent, and the new limits were

$$p' \pm 3\sqrt{\frac{p'(1 - p')}{n}} = 1.2\% \pm 3\sqrt{\frac{1.2 \times 98.8}{200}}\% = 1.2\% \pm 2.3\%$$

The limits therefore were 3.5 per cent and 0 per cent. In terms of the allowable number of defects per sample, each sample of 200 was permitted to contain between 0 and 7 defective covers.

Sampling inspection was possible in this case because, (*a*) the chart had established a control on the plating work that held the subcontractor at a controlled quality level; (*b*) as long as the plating was not more than 3½ per cent bad, it was cheaper to throw out the bad plating at assembly than to inspect it 100 per cent at receiving; and (*c*) if an excessively bad lot should appear, the sample was practically as effective as 100 per cent inspection in rejecting it.

Using this chart resulted in a saving of inspection man power by the prime contractor and an increase in acceptable work by the subcontractor.

CASE HISTORY V. *pn* CHART, COMPARISON OF THREE VENDORS

A prime contractor had subcontracted to three different shops the manufacture of a metal part. Vendor 1 machined the part from castings, vendor 2 used bar stock, and vendor 3 used forgings. Each had accepted the contract at a different price, and, as Chart V shows, each had his own "pattern" of bad work. The summary data are given in Table 14.

Chart V was drawn from the prime contractor's receiving-inspection records. From each lot that came in, a sample of 300 was taken, and the number of rejects was recorded. The respective averages and limits for each vendor appear on the chart. (The highest point on each chart was omitted, because they all occurred on the same day, when the chief inspector was absent and his assistant was unnecessarily strict in inspecting.)

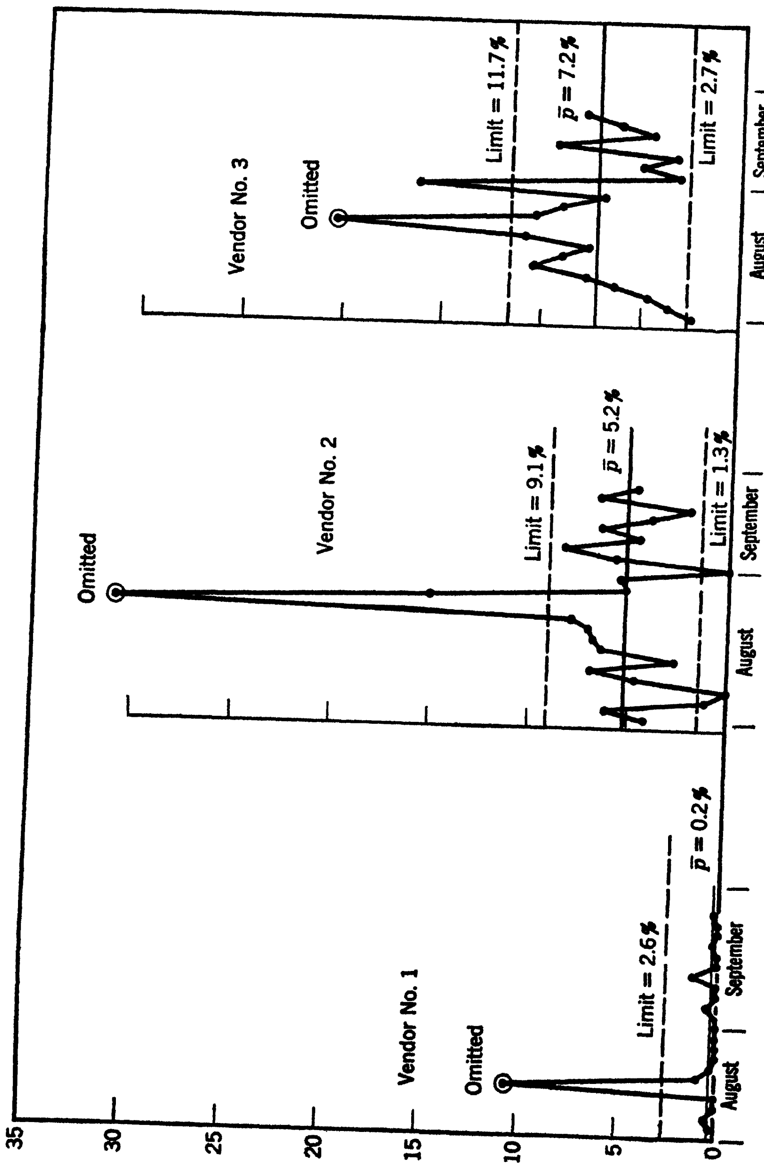


Chart V. Comparison of three vendors

Of the three vendors there was no question which was best. Not only did vendor 1 have the lowest $\bar{p}$, he was also the only one exhibiting good control: his work was the most dependable. Interestingly enough, he was also the lowest bidder. An investigation revealed that the process more than the vendor accounted for the superiority. Thereafter all these parts were made from castings, with a considerable reduction in cost as the result.

TABLE 14

	Raw Material	Cost per Piece to Prime Contractor	Average Per Cent Rejected at Prime's Receiving Inspection
Vendor 1	Castings	72¢	0.2
Vendor 2	Bar stock	84¢	5.2
Vendor 3	Forgings	82¢	7.2

Sample inspection and control charts used in the receiving department make possible a scientific evaluation of vendors' quality. They enable acceptance or rejection decisions to be made in a scientific and economical way.

CASE HISTORY VI. COMPARISON OF THREAD GAUGES BY PRIME AND SUBCONTRACTOR

A major airframe manufacturer let a subcontract to an aircraft-parts supplier for a large number of metal tubes threaded with pipe thread on one end. The subcontractor's production proved to be unsatisfactory—up to 50 per cent of his output failing to meet gauge specifications at the prime contractor's receiving inspection. Mutual recriminations soon built up ill will, the vendor feeling that the prime contractor was unduly harsh in his standards and the airframe manufacturer insisting that specifications had to be rigidly enforced. Relations between the two men were going from bad to worse when one of the prime contractor's men who had studied statistical methods of quality control introduced the subcontractor to control-chart techniques.

First, they agreed what deviations were acceptable. It was decided that a deviation from fit of not more than three threads on the thread

gauge would be permissible. Next, the subcontractor put an inspector at each of his machines; written inspection records were kept and put on control charts in order to find out what was happening in the production process. The subcontractor then put his engineers to work tracing down the reasons for each out-of-control point. Their investigations eliminated the problems one by one until the process was fairly well in control and rejections had dropped from around 50 per cent to 2 or 3 per cent. Thereafter the subcontractor reduced his process inspection to periodic small samples, depending on his control charts to warn him of impending trouble.

At this point the two parties agreed to maintain identical control charts for the purpose of acceptance or rejection of deliveries. The tubes were packed in boxes of 100, ten boxes to a carton, so that each carton contained 1,000 tubes. Both prime and subcontractors inspected one tube from every two boxes, making a sample of five bolts from each carton, the inspections being independent of each other. With each shipment the subcontractor sent a copy of his control chart, based upon the average deviations from fit of the five sample tubes. The allowable individual deviation being three threads on the thread gauge, the deviation of an average of five would be $3/\sqrt{n} = 3/\sqrt{5} = 1.345$. Chart VI shows the two control charts—the vendor's and the purchaser's—superimposed, with the standard limits set at ± 1.345. In the lower part of Chart VI is shown the difference between the vendor's and the purchaser's samples, carton by carton.

The points marked 1 and 2 on the chart indicate two cartons that the airframe manufacturer rejected. Both charts clearly proved the unacceptability of these lots of tubes, and the subcontractor therefore willingly agreed to take them back and detail them. A few days later, at point 3 the vendor objected to a rejection, insisting that his chart evidenced good quality although the prime contractor's chart showed out of control. Again at point 4 the charts failed to agree. This time the subcontractor refused to take back the cartons that were assertedly substandard. Both parties then agreed at 5 to an investigation, part of which was the preparation of the comparison chart. Here a clue was found in the fact that beginning with the 24th carton 17 successive difference points were below the zero line. On the theory of runs (see page 64), this indicated an assignable cause of difference between the vendor's and purchaser's inspections, with the vendor's record consistently lower than the purchaser's. A brief investigation led to checking the gauges of both disputants: the prime contractor's gauge proved to be the faulty one.

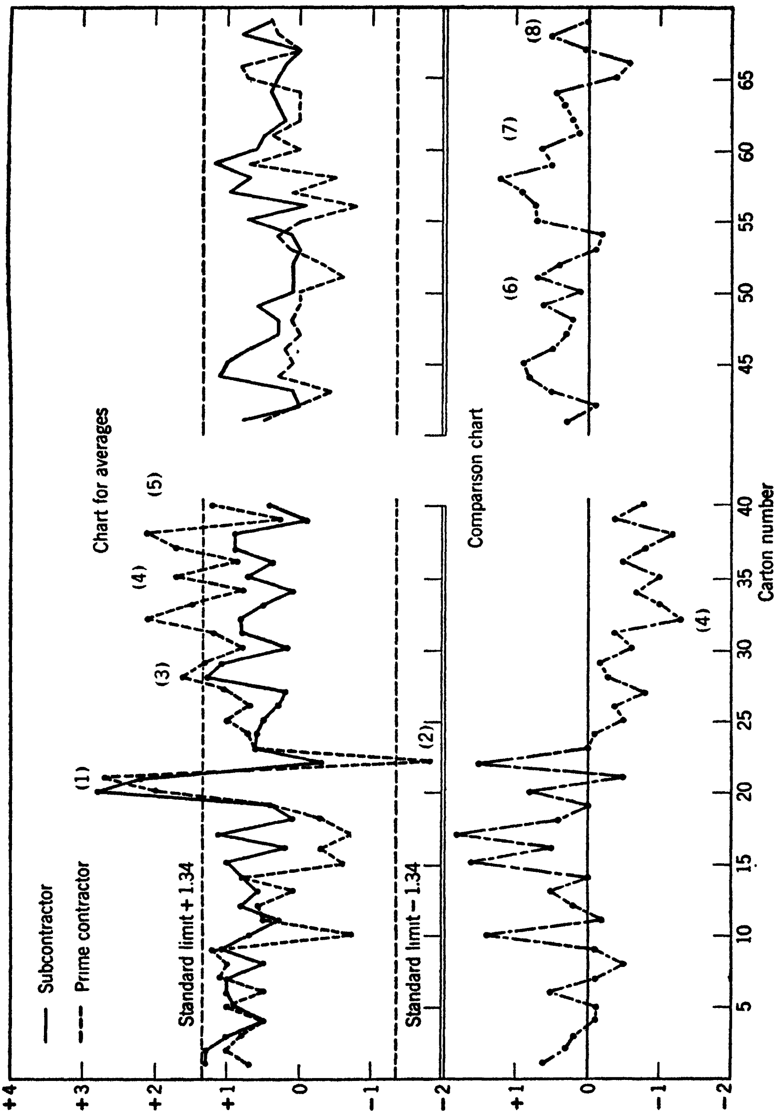


Chart VI. Comparison of thread gauges by prime and subcontractor

After a good gauge had been substituted for the bad one, cartons 41 to 69 appeared as shown on the right-hand side of Chart VI. All points were now in control, but the difference chart showed 22 points out of 29 *above* the zero line: at point 6, ten successive points above zero; at 7, another ten above zero. All these indicated a discrepancy, but not a large enough discrepancy to result in lack of control or disagreement about quality. The situation, however, bore watching, because it was properly interpreted as a tendency for the subcontractor's gauge to become faulty. So helpful were these charts to both parties that they were continued during the life of the contract.

Not always can disputes between seller and buyer, or even between one department and another within the same company, be settled as easily as this one was. Very often such disagreements about quality create ill will and misunderstanding. On the other hand, a meeting of the minds as to acceptable quality, coupled with the collection and analysis of appropriate facts, can avoid costly arguments and unnecessary expense.

CASE HISTORY VII. A PROBLEM IN SUBGROUPING

One of the component parts of a certain machine tool was an eccentric cam in which a slot one-half inch deep was milled. The operator of the milling machine placed five cams in the jig, tightened them in with an adjusting screw, and cut five slots simultaneously.

During a study of this milling operation, it became apparent that the slots were not being held within tolerances. Both foreman and

TABLE 15

Sample No.	Individual Measurements of Five Slot Depths, Inch					Average	Range
	a	*b*	*c*	*d*	*e*		
1	0.4990	0.4995	0.5000	0.5000	0.4995	0.4996	0.0010
2	0.4995	0.4990	0.4995	0.4995	0.4995	0.4994	0.0005
3	0.5000	0.5000	0.5005	0.4995	0.5000	0.5000	0.0010
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

$\bar{\bar{X}} = 0.4989$

$\bar{R} = 0.0008$

operator complained that the milling cutter had to be changed too often and that the slot went out of tolerance before the cutter needed resharpening. In order to reduce the amount of down time required for changing cutters and to increase the efficiency of the operation, a control-chart analysis was made.

Since five slots were cut at one time, it seemed logical to use groups of five simultaneous slots as the subgroup for the control chart. Measurements of slot depth therefore were made on each of five simultaneous slots, about half an hour elapsing between inspections. The original data were set up in the form shown in Table 15.

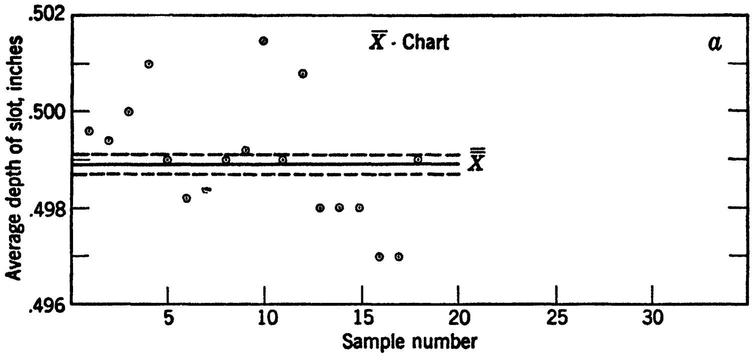


Chart VIIa. Milling slots in cams. Samples of 5

Chart VIIa shows the $\bar{X}$ chart for 18 samples of 5 each, together with control limits calculated from the $\bar{\bar{X}}$ and $\bar{R}$ values already given. Either the process was very erratic or the subgrouping was wrong. A classification of the sources of variation in the operation developed the following information:

1. Raw materials—variability in hardness of steel.
2. Dimensions of cam—variability from preceding operations.
3. Positioning of the milling jig—variability due to operator skill.
4. Wear in the cutting tool—variability caused by slot growing shallower as cutter became dull.

So far as the chosen method of subgrouping was concerned, the cams were thoroughly scrambled or randomized before coming to the mill. Causes 1 and 2 therefore were *included* in the sample, because a cam of any specified hardness or size would be just as likely to appear in one sample as another.

Causes 3 and 4, however, were *not* included in the sample, but had their effect *between* the samples. Positioning of the jig affected all the five cams in the jig in the same way, but the next group of five might

be positioned differently. Tool wear, being a long-term directional effect, should appear as a trend in successive samples and not as a range within samples. Regarding cause 3—positioning of the jig—some variation between successive jig settings was unavoidable, since jig setting was dependent upon the manual skill of the operator; certainly *excessive* variations from this cause were undesirable.

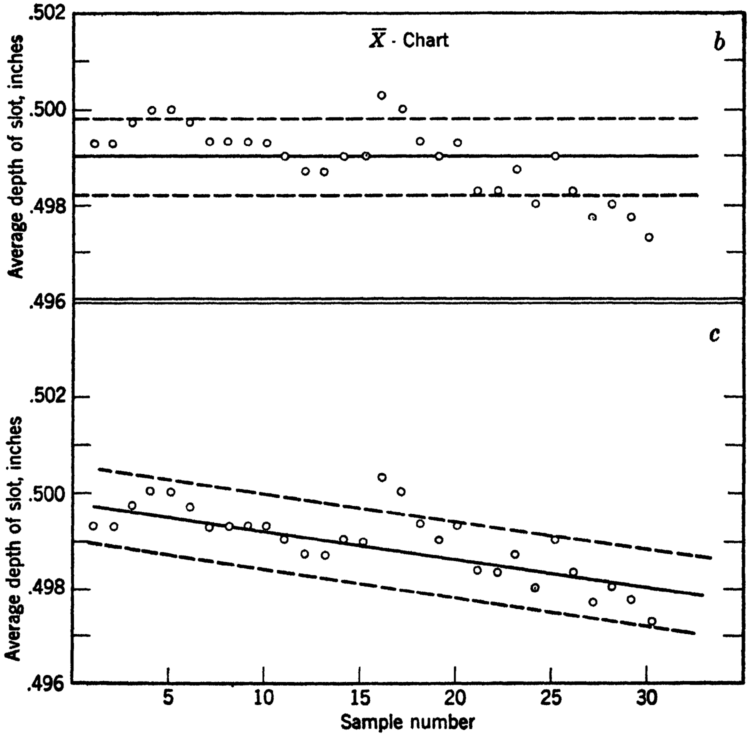


Chart VIIb. Milling slots in cams. Samples of 3

Chart VIIc. Milling slots in cams. Samples of 3 with trend

With these considerations in mind, it was decided to *include* jig positioning in the within-sample variation by changing the subgroup method. The inspector was told to take one cam off each of three successive cuts and to repeat the procedure every half-hour. This would, in effect, include normal variations in hardness, dimensions, and positioning within the sample, leaving only tool wear outside. Chart VIIb shows the results with samples of three taken by the new method. The trend of tool wear, faintly visible in Chart VIIa, then became clear, leading to the conclusion that jig setting had been a

major cause of variation, so great as to have masked the tool-wear trend in Chart (a). Samples 16 and 17 are out of control on Chart (b), indicating at those points an undue carelessness on the part of the operator in setting the jig.

Chart (b) also shows the samples at the top and the bottom of the trend to be out of control. A least-squares trend line therefore was calculated,[1] and limit lines were drawn parallel to the trend at a distance $\pm A_2\bar{R}$ from it along the perpendicular axis. Samples 16 and 17 were still out of control, as was sample 30. A recalculation of the trend, excluding samples 16 and 17 from the calculations, gave the picture shown on Chart VIIc. Here the trend of tool wear had been isolated and measured, and limits had been set for normal variations due to other causes. An out-of-control point therefore represents a truly assignable cause, one that should be hunted for and eliminated. Thus Chart (c) fulfills the basic function of a control chart, which is to discover *those faults and errors that are controllable but are not controlled.*

Neither Chart (a) nor Chart (b) accomplished this basic purpose; on both of them the out-of-control points represented to some extent causes that were an integral part of the operation and could not be eliminated. This study illustrates how little value a control chart has unless, in operationally verifiable terms, it points to economic improvements or to predictable stability in the manufacturing process.

CASE HISTORY VIII. GOOD AND BAD OPERATIONS

In small plants as well as large ones, control charts have been used successfully. To the small plant a good operation, one in control at a satisfactory level, is often more necessary than to a large plant. When a company with only a few machines takes on a single contract absorbing its entire capacity, or does job-lot business in short runs on a variety of products, the successful operation of the equipment it has is of paramount importance. With all its eggs in one or a few baskets, and with its dependence upon other manufacturers for business, the small company cannot afford to take chances with its reputation for quality and low cost.

A certain machine shop had two automatic screw machines: a fairly new six-spindle and a very old four-spindle. One employee, as combination foreman–setup-man–operator, ran both machines. Neither machine was performing as it should, but the old one especially required a constant struggle to keep it going. The operator spent three

[1] See any standard statistical text for the method of calculation.

quarters of his time adjusting and readjusting the machine, changing and resharpening the tools. Efficiency was less than 60 per cent. The company was unable to meet scheduled requirements. Taken to task by his boss, the operator said, "What can you expect out of an old rattletrap like that?"

Chart VIII pictures the results obtained from a statistical study of the old machine on two complete runs of a steel turning. Chart (a) shows the original condition of the operation and Chart (b) the improvement that was made. Several points are worthy of note in these two charts.

Samples of four successive pieces, one off each spindle, were taken and measured with a dial indicator to the nearest 0.0001 inch. From these samples the $\bar{X}$ and R charts were drawn. Chart (a) illustrates a typical badly out-of-control operation: the $\bar{X}$ chart points out difficulties in tool setting and adjusting; the R chart indicates assignable causes of trouble within the machine itself. Together they reveal that not all the problems were due to the machine alone; if that had been true, the $\bar{X}$ chart would not have been so unpredictable as it was. Judging from the violent fluctuations in the $\bar{X}$ chart, the operator was unable to estimate his setup accurately. This difficulty was enhanced by the unusual lack of control in the range chart, where no fewer than 9 out of 78 points fell beyond the $D_4\bar{R}$ limit. It is interesting to note that several times the R chart went out of control just before or during trouble with the $\bar{X}$ chart. Several significant points are indicated on Chart (a) by numerals in parentheses. They refer to the following comments:

1. Samples 1–5. A poor setup, with an average too low. Fault of the machine, as shown by out-of-control ranges. One of the spindles went bad. The single point at sample 6 on the $\bar{X}$ chart represents an unsatisfactory experimental setup made during efforts to correct the bad spindle.

2. Sample 18. Tools resharpened too late. Should have been done two samples earlier. The out-of-control range at sample 16 resulted in an inaccurate estimate of the next two sample averages because of erratic behavior of one of the spindles. The machine was shut down for repairs.

3. Samples 19–30. After being repaired, the machine performed fairly well, except for out-of-control ranges at samples 24 and 25, until sample 30. Here the operator had left the machine running while he had lunch. Sample 30 was taken during the lunch hour and was found to be very much too high on the $\bar{X}$ chart. The new setup made at sample 31 was excellent, starting at the lower control limit.

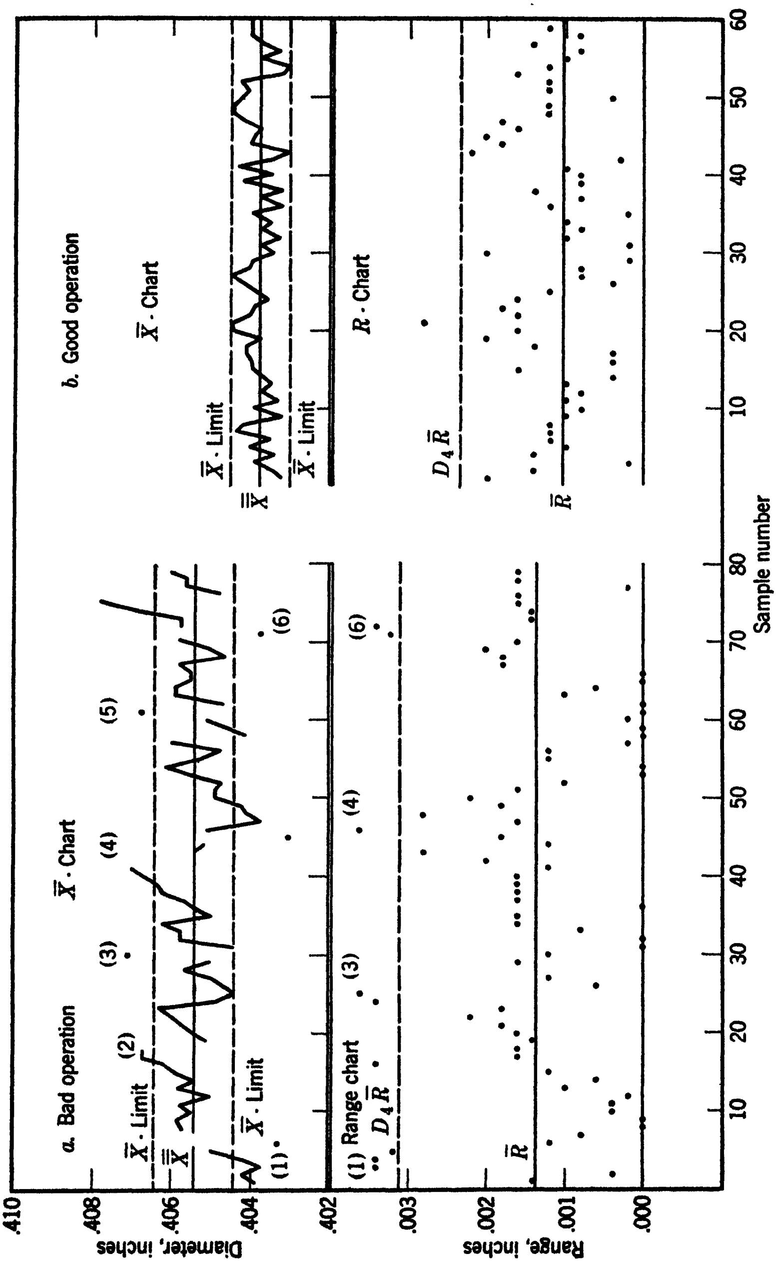


Chart VIII. Good and bad operations. Four-spindle automatic screw machine. Samples of 4: one piece off each spindle

4. Samples 41–47. A four-hour struggle to get the machine operating satisfactorily. The group of high range points (with one out of control) indicated more mechanical trouble. There was practically no production during this period.

5. Samples 58–61. On sample 58 the shop inspector stopped the machine because his spot inspection disclosed an out-of-tolerance piece. Blaming it on the machine, the operator spent most of the next two hours trying to find a probably nonexistent trouble. During this time the ranges were below the $\bar{R}$, indicating consistent work by all spindles. When at last the operator discovered a soft spot in the steel bar, he was satisfied and went back to production. (Later studies showed that within wide limits of Rockwell tests, hardness had no significant effect on the dimension of this part.)

6. Samples 71–75. More spindle trouble and some operator carelessness. Keeping of this chart and frequent questioning of the operator had produced in him a psychological resistance which made further analysis difficult. Since the run by this time had been completed and adequate data had been gathered, the chart was temporarily discontinued.

The conclusions reached and action taken as the result of Chart (a) were:

1. A large part of the trouble was caused by mechanical defects in the machine. Since a new one could not be bought under war conditions, the old four-spindle automatic was shut down for a complete overhaul and repair job.

2. A fair share of the difficulty was caused by inaccurate setups and by sharpening and changing tools at the wrong time. This was due to the operator's one-piece inspection procedure. His practice was to measure one piece off the machine, and to start, adjust, or stop the operation accordingly. At best this would result in frequent misadjustments, because measurement of one piece cannot give adequate information about this kind of process. The operator, for instance, had no idea what was the average size of the parts he produced. When the $\bar{X}$ chart revealed that the $\bar{\bar{X}}$ was 0.4054 inch, he was incredulous. He had been aiming at "about 0.404 inch." He had been shown the chart as it developed and frequently had objected to the pointed questions it raised.

To correct the operator's inaccuracy in positioning the tools, he was instructed to measure one piece off each spindle before he started the machine on production after a shutdown, no matter how short or long the shutdown might be. If the average and range in the sample of four pieces failed to meet the standards set by management, he was to

take another sample of four pieces and continue setup or other adjustments until two successive four-piece samples met the standards. Not until then could he start the machine on its run.

The standards differed from product to product, depending upon specifications and operating tolerances. For the product discussed here, Chart VIIIb shows the result obtained with the repaired machine and with more accurate tool setting. Chart (b) successfully met management standards except for sample 21 on the range chart where a brief flare-up of spindle trouble was brought quickly under control.

An interesting feature of the (a) and (b) range charts is the presence of 13 zero ranges on (a) (where they would not be expected), and of no zeros on (b). One by-product of the investigation made in this case was the discovery of a defect in the dial indicator used. This had been corrected before the (b) range chart was made.

In Table 16 are given the $\bar{\bar{X}}$, $\bar{R}$, and limits for both bad and good operations.

TABLE 16

	Bad Operation (1/1,000 Inch)	Good Operation (1/1,000 Inch)
$\bar{\bar{X}}$	405.38	403.75
$\bar{R}$	1.37	1.03
A_2	0.729	0.729
$A_2\bar{R}$	1.00	0.75
I_2	1.46	1.46
$I_2\bar{R}$	2.00	1.50
$\bar{\bar{X}} + A_2\bar{R}$	406.38	404.50
$\bar{\bar{X}} - A_2\bar{R}$	404.38	403.00
$\bar{\bar{X}} + I_2\bar{R}$	407.38	405.25
$\bar{\bar{X}} - I_2\bar{R}$	403.38	402.25
D_4	2.282	2.282
$D_4\bar{R}$	3.12	2.35
D_3	0	0
$D_3\bar{R}$	0	0

CASE HISTORY IX. THERMOSTAT CONTROL

A manufacturer of industrial heat-control equipment produced a thermostat with a specified operating range of 500 to 600 degrees Fahrenheit. It was designed to turn on an electric current if the temperature fell below 500 degrees and to break the circuit if it rose above 600. Production was about 1,000 a day on two assembly lines. The

final assembly operation was performed by one girl on each line who made the final adjustment of the thermostat. A sample from each day's production was given a rigid test in one of two special testing machines which simulated the conditions the thermostats would meet in actual use. Each of the machines was able to test 25 thermostats in a ten-hour shift. With the plant working two shifts this permitted not more than 100 tests per day to be made. One hundred per cent testing was therefore impossible. In order to give himself assurance that the untested product was meeting specifications, the manufacturer decided to use control charts on his test results. Charts IXa, b, and c describe his experience in collecting his original data and in analyzing them for assignable causes.

Measurements were made of the individual operating temperature of each of the tested thermostats. For convenience they were taken in five groups of five off each machine on each shift. Identification tags were placed on each thermostat showing the test machine, the initial of the girl who made the final assembly adjustment, and the shift on which it was assembled. Averages and ranges were calculated and plotted each day for 20 samples of five measurements each. Table 17 gives a typical day's record from which Chart IXa was prepared.

Ninety samples of five measurements each were collected over a five-day period. Averages and limits were calculated for Chart IXa as follows:

$$\begin{aligned}
\bar{\bar{X}} &= 530 \\
\bar{R} &= 38.6 \\
\bar{\bar{X}} \pm A_2\bar{R} &= 552,\ 508 \\
D_3\bar{R} &= 0 \\
D_4\bar{R} &= 81.6
\end{aligned}$$

Both the average and the range were out of control. Apparent during the first three days was a downward drift in the averages. Of the first day's 20 samples, only one was below the $\bar{\bar{X}}$ line; a run of 12 consecutive points was above the average. The second day was fairly well centered. On the third day 13 consecutive points were below the average. Not until the fifth day did the $\bar{X}$'s again become grouped normally about the $\bar{\bar{X}}$ line. Suspicion immediately arose that an almost imperceptible and progressive change was taking place in the settings of the test machines.

To determine the cause of the drift Chart IXb was drawn, grouping all the A-machine tests together chronologically and all the B tests similarly. On machine A the drift was more marked than on machine B. A closer examination showed that the first day was above average on both machines; if that day were removed from the data, machine A

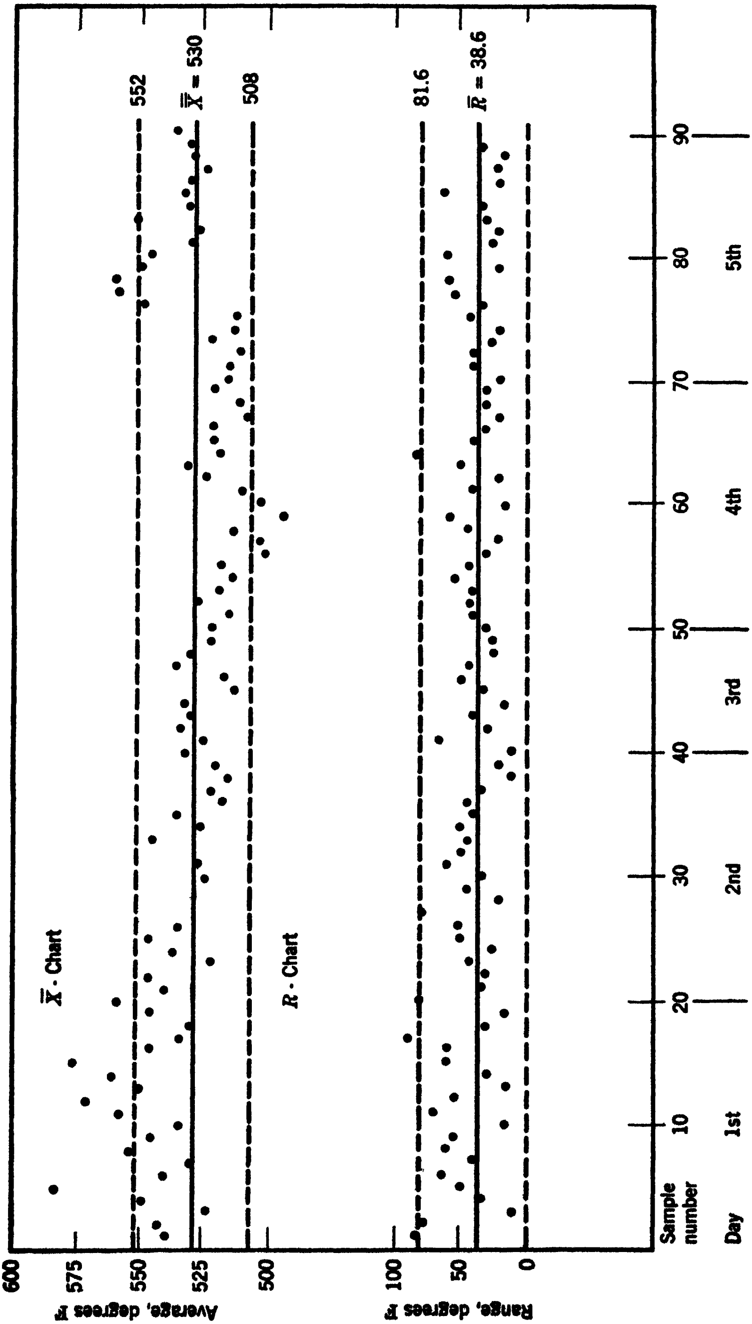


Chart IXa. Thermostat Control. By days. Samples of 5

TABLE 17

Date	Sample No.	Testing Machine	Shift	Assembly Adjustment by	Average, Degrees ($n = 5$) $\bar{X}$	Range, Degrees R
Jan 8	1	*A*	1	N	541	85 *
	2	*A*	1	N	544	80
	3	*A*	1	N	525	10
	4	*A*	1	N	551	35
	5	*A*	1	N	584 *	50
	6	*B*	1	D	542	65
	7	*B*	1	D	531	40
	8	*B*	1	D	555 *	60
	9	*B*	1	D	546	55
	10	*B*	1	D	536	15
	11	*A*	2	B	559 *	70
	12	*A*	2	B	572 *	55
	13	*A*	2	B	551	15
	14	*A*	2	B	562 *	30
	15	*A*	2	B	577 *	60
	16	*B*	2	P	547	60
	17	*B*	2	P	535	90 *
	18	*B*	2	P	531	30
	19	*B*	2	P	547	15
	20	*B*	2	P	560 *	80

* Out of control

would show a less obvious drift while machine *B* would show none at all. Clearly some assignable cause had operated on the first day. Beyond that, machine *A* showed a somewhat biased trend.

Seven sample averages were out of control on the first day:

5 on machine *A*, 2 on machine *B*,
5 on the second shift, 2 on the first shift,
4 on operator *B*, 1 on *N*, 1 on *D*, 1 on *P*.

Operator B was a new assembly adjuster who had not yet acquired the necessary skill. Taking her errors out left one or two out-of-control points on each testing machine, on each shift, and on each operator. This pointed to a flaw affecting all parts of the process, probably raw material. It was later discovered that the bimetal used on that day was at fault.

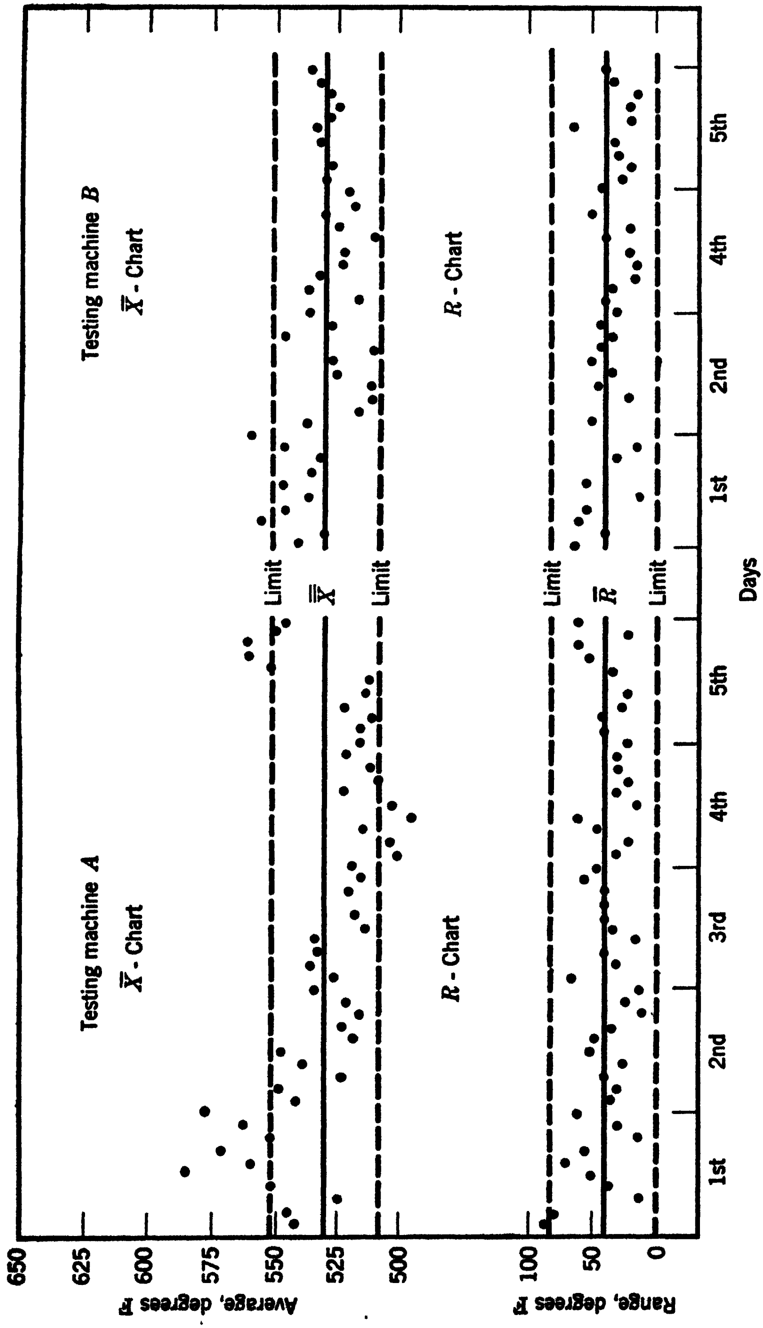


Chart IXb. Analysis by testing machines. Samples of 5

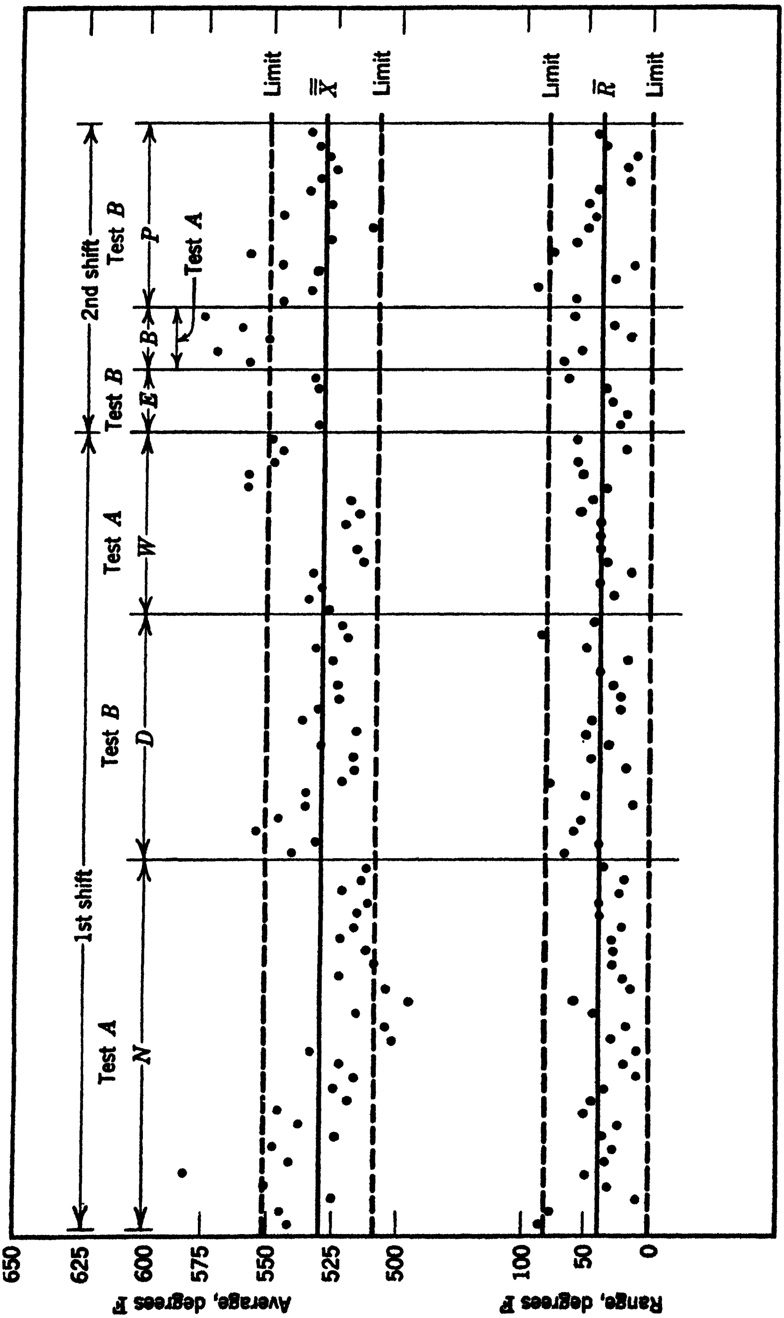


Chart IXc. Analysis by operators and shifts. Samples of 5

Chart IXc shows an analysis by adjusters and by shift. Taking out, in imagination, the first day's record, and examining each adjuster's chart for runs above and below average, adjuster N showed up as definitely an assignable cause, while D and W were suspects. The whole first shift seemed below average. Much of the drift seen in Chart IXa was traceable to an apparently gradual change in adjuster N's ideas of correct temperature. She seemed to favor the low side. Adjuster D was fairly well centered though somewhat low. Adjuster W changed her mind suddenly on the fourth day. Adjusters E and P on the second shift were accurate and reliable. Adjuster B, after her first-day attempt, was not used again on this critical and highly skilled job.

On the range charts three points were out of control. Two of these, on the first day, were attributable to defective bimetal; the third came from adjuster D, a highly skilled girl and was never traced to any particular cause—it was apparently an accidental variation.

As analyzed on these charts, the process was not satisfactory. The dispersion of individual observations as estimated by taking $2I_2\bar{R} = 2 \times 1.15 \times 38.6 = 88.8$ degrees was within the tolerance range of 100 degrees, but the average of 530 degrees was too low. Work began, as soon as the analysis had been completed, on two programs. First, an engineering study was begun to find the most economical way of raising the process average to 550 degrees; second, control charts were set up as a permanent inspection procedure for assuring a consistently good product.

CASE HISTORY X. NECESSITY FOR $\bar{X}$ AND R CHARTS

Chart X tells graphically why the calculation of averages and dispersion is necessary in order to discover causes of variation in a manufacturing process. The operator, foreman, or inspector who inspects one piece in order to decide whether or not the process is satisfactory, may happen to inspect any one of the pieces represented by the points on Chart Xa. How can he tell what the process is like unless he knows where that one measurement is relative to the others? If, on the other hand, he measures several pieces and performs the simple arithmetic necessary to calculate their *average*, he knows within certain limits what the *average of the process is*. If he keeps a graphic record of each average in the order of production, he begins to see trends, tool wear, setup positions, and other features of the process as they are revealed in Chart IXb. Such records show, of course, only measurements made on certain selected pieces; but with a substantial quantity of

known measurements recorded in this way it is possible to diagnose the condition of the process as a whole, that is, of the pieces *not measured.* People tend to forget that what they see is not the whole truth, perhaps not even a partial truth.

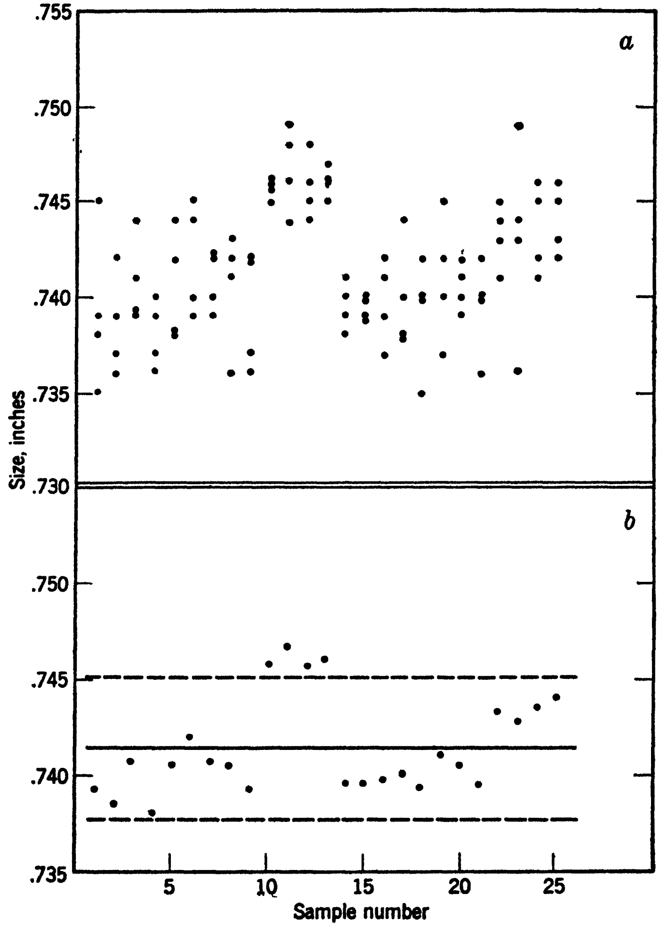


Chart X. Individual versus average charts

(a) Chart for individual observations. Metal knobs of Figure 1a from Table 3
(b) Chart for averages of 4

What has been said of averages holds equally well for dispersion. A sample of one observation gives no idea whatsoever whether the process is holding specification limits. Even a sample of several pieces is inadequate for that purpose. The only way in which a correct picture of the

variation can be obtained is by some estimate of the dispersion from a single large sample or preferably from a *sequence of small samples.* This is the secret and power of the control chart: it provides an easy, sensitive, and satisfactory measure of process variation, by accumulating small samples into large ones. For a controlled process the dispersion of individual measurements is approximately $2I_2\overline{R}$; for a series of small samples of size n it is $D_4\overline{R} - D_3\overline{R}$. As with averages, so with dispersion: an inadequate *quantity* of inspection is bad enough, but even an adequate quantity of inspection fails to achieve its purpose unless two further steps are taken. First, record the data; second, analyze it.

Charts Xa and b are particularly illuminating in this respect. Could any man, no matter how intelligent, get from all the separate measurements of Chart (a), taken over a period of days, the kind of picture shown in Chart (b)—the comprehensive, progressive summary of a process life, and the certainty of knowledge which it gives?

CASE HISTORY XI. GOOD AND BAD WORKMANSHIP

At one stage in the manufacture of a certain article, a small part was shaped from sheet steel in an automatic stamping press. The metal had previously been cut in an irregular shape to fit the stamping operation. The precut pieces had to be inserted in the stamping press in a particular way in order to be formed satisfactorily. A good deal of manual dexterity was required of the girl at the stamping press, since the operation was a rapid one—about 1,000 stampings per hour—and the insertion was made by hand. Occasionally even the best operator would fail to insert the piece in just the right way: a defective article would be the result. Process inspection consisted of the inspector examining 20 pieces per hour and recording the number of defective stampings which appeared in his sample.

In Chart XI is shown the record of two operators over a six-month period. Each point on the chart represents the results of a week's inspection, that is 20 pieces per hour for 50 hours. N therefore equals 1,000. Below is given the average per cent of bad work and the normally expected limits of variation for each of the two girls, based upon the formula $\bar{p} \pm 3\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$.

	Operator B	Operator A
Best expected quality	0.2% defective	4.7% defective
Average quality	1.3% defective	7.1% defective
Worst expected quality	2.4% defective	9.5% defective

Operator A was in control, but at a high and unsatisfactory level. No particular week or series of weeks could be considered significantly worse than any other. Her problem was the basic one of not having acquired the necessary skill, or of not being adapted to this type of work. Occasional reprimands would do no good, because, though her

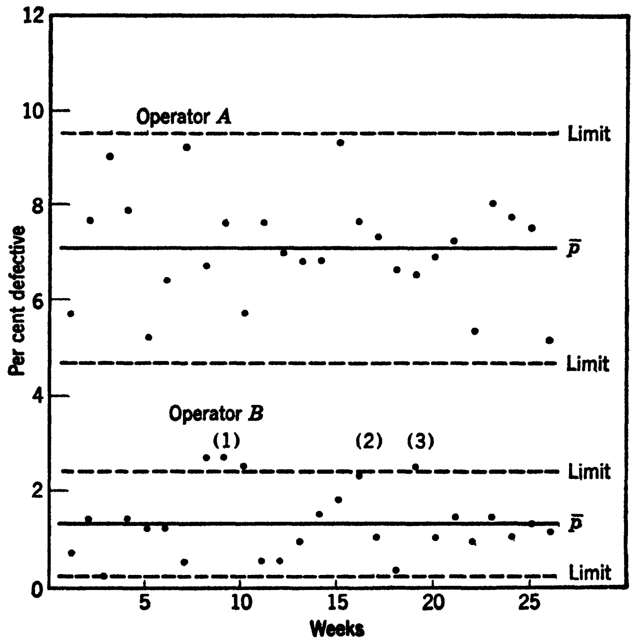


Chart XI. Good and bad workmanship. Comparison of two operators. Per cent rejections by weeks

work was poor, no specific week was significantly worse than the week before. A reprimand, therefore, would have no clearly defined fault of the operator to justify it.

Operator B, on the contrary, shows evidence of good though erratic work. At (1) on Chart XI operator B took a three-week leave of absence; a substitute took her place with the result that relatively poor work was performed all during her absence.

After the regular operator returned to work she did very well for the first two weeks. Thereafter for a month her work grew progressively worse, stopping just short of the upper limit in the 16th week. The record showed that a department head had taken her to task for carelessness during the 15th week; it is doubtful whether he should have

done so, for such a progression of points could easily occur by chance, especially since the chart did not go out of control.

At (3) the story was different. Operator B should have been asked at that point why her work was not up to standard; such questioning, sympathetically done, might have revealed some assignable cause of trouble that could have been eliminated. After this chart had been drawn from the inspection records, the girl was questioned and readily admitted that a serious family problem had arisen at the time she took her leave and that at about the time of the last out-of-control point it had finally been solved. Thereafter, as the chart shows, she settled down to normal again.

This kind of personnel record enables workers to be divided into two categories, each of which requires a different approach. Those, like operator A, who are *consistently poor* need constant supervision and careful training. Those, like operator B, who are good but erratic need occasional attention, especially when their work is worse than could happen by chance. If each worker is given the right kind of attention, at the right time, quality can be steadily improved, both for individual persons and for the plant as a whole.

CASE HISTORY XII. HAMMER-OPERATOR QUALITY

Chart XII illustrates several useful features of control charts. The form in which the original data are combined with the chart itself makes for economy in keeping records and convenience in studying them. Sample sizes as small as 10 observations can be used effectively; and the explanations of excessively high points, representing unusually poor work, gives a permanent, growing mass of information as to causes of trouble.

In a commercial forging plant, doing custom forging with relatively short runs on a large variety of products, a chart of this sort was kept on each hammer operator. The man whose record is shown was one of the best, an experienced hammer man and an old employee. Each of the other operators had his own chart, from which monthly quality ratings were calculated and which were used as a basis for individual study of each employee aimed at constant improvement of his work.

Process inspection consisted of periodic sampling, with sample size dependent upon the rate of production. Each day the percentage of bad work was determined for each job worked on by each operator. The first column in Chart XII gives the date; the next, the machine or hammer number; the third, the job or part number. The next three

columns show n, the number of pieces inspected on that day at that job; c, the number of bad pieces found; p, the per cent, that is $c/n \cdot 100$.

Date	Mach.	Part	n	c	p	Insp.
Mar. 1	1	X·3	34	0	0	A
2	1	X·4	47	0	0	A
5	1	X·3	48	0	0	A
6	1	X·3	68	1	1.5	A
7	1	X·3	55	0	0	A
8	1	X 3	42	0	0	A
12	1	X·4	127	0	0	A
13	1	X·4	118	0	0	A
14	1	N·2	20	1	10.0	A
15	1	X·2	52	0	0	A
16	1	N·2	27	0	0	A
16	1	X·4	14	0	0	A
19	1	X·10	12	0	0	A
19	1	X·9	52	0	0	A
20	1	X·10	54	0	0	A
21	1	N·2	35	0	0	A
21	1	X·10	17	0	0	A
21	5	X·3	31	1	3.0	A
22	5	X·4	67	3	4.5	A
23	7	X·3	72	2	2.8	A
24	7	X·3	53	4	7.6	A
24	5	X·4	28	2	7.1	A
27	5	X·3	133	3	2.3	B
28	5	X·3	85	2	2.4	B
30	5	X·2	65	3	4.6	B
Apr. 2	1	X·10	56	0	0	B
3	1	X·4	121	1	0.8	B
4	5	X·2	19	1	5.3	B
4	1	X·4	9	0	0	B
5	1	X·4	54	0	0	B
6	1	X·4	56	0	0	B
10	1	X·3	53	1	1.9	B
11	1	X·3	53	0	0	B
12	1	X·3	45	0	0	B
13	1	X·3	45	0	0	B
14	1	X·3	38	1	2.6	B
16	1	X·3	25	0	0	B
17	1	X·3	37	0	0	B
18	1	X·3	13	0	0	B
19	1	X·8	62	4	6.4	B
22	1	X·10	24	0	0	B
25	1	X·1	48	0	0	A
26	1	X·1	48	0	0	A
27	1	X·1	54	0	0	A
28	1	X·1	6	0	0	A

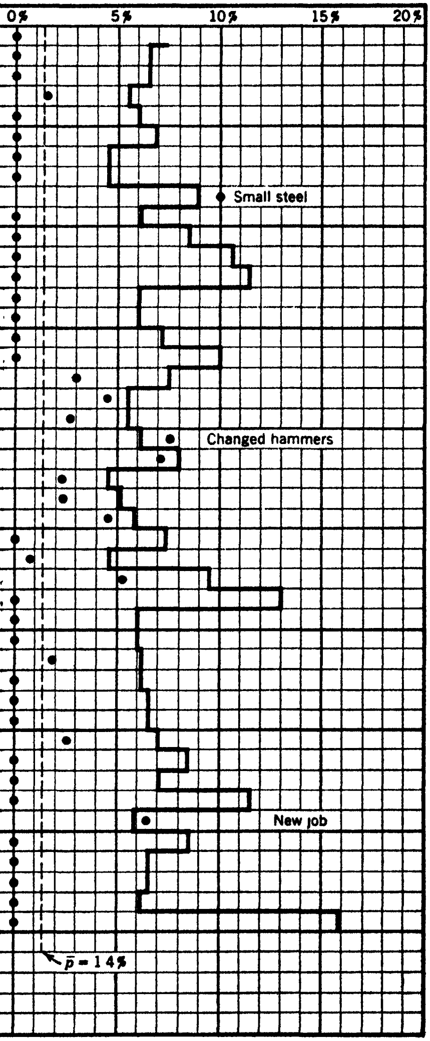


Chart XII. Hammer operator quality

"Insp." designates the inspector who collected the data. On the chart itself the figure in the p column was plotted; $\bar{p}$ was calculated and drawn in; and $\pm 3\sigma_p$ control limits were set up for each point, depending upon the number of pieces inspected, and derived from Figure 5

in Chapter 4. This procedure resulted in a permanent personnel record of great value: it discovered ways to help each man increase his skill.

Studying this and other charts led the shop superintendent to realize his two principal sources of bad work. First, a man did better if kept on the same hammer as much as possible; second, he did better if kept on the same kind of forgings as much as possible. These two conclusions led to a radical change in scheduling policy. Each operator was assigned a hammer of his own, and each hammer was scheduled on jobs as similar as possible. Desirable jobs and desirable hammers were made the reward in quality competitions. This shop is now producing forgings to strict standards with an average of less than 2 per cent unacceptable work, including both repairs and scrap—an excellent record under usual commercial conditions.

A valuable by-product of these charts was their ability to diagnose an operator's work habits. One man, for instance, usually started off a new job with a high percentage of bad work; he took three or four days in settling down to his normal performance. Observation of and conference with this man revealed him to be a conscientious, hard worker with a nervous temperament. Thereafter, whenever he was assigned to a new job the foreman would drop by and say, "Now, Bill, take it easy," relieving his tension and encouraging him to more self-confidence. Another operator had a sick wife with whom he had to stay up at night; his work suffered in consequence. When his chart showed a serious upward trend in poor work the foreman pointed it out and asked him why. When he had explained his family problem, the foreman told him about the company's medical insurance plan which provided hospitalization at a nominal cost. After the man had put his wife in the hospital his work immediately improved to his normal level.

A record such as this makes it possible to visualize the progress made by an apprentice or a new employee, and to weed out the unpromising ones before too much money has been spent in trying to train them. It encourages pride of workmanship. It provides a basis for promotion entirely free from accusations of favoritism. It leads to improvement of equipment and production methods. Not only in forging but also in many other manufacturing operations where operator skill is paramount, personnel control charts have proved their value as tools of management.

CASE HISTORY XIII. PRODUCTION VOLUME OF INDIVIDUAL WORKERS

In a certain plant operating with an incentive pay plan, standard production quotas had been set for each kind of job by means of time studies. Production workers earned extra pay for production over and above the standard task. Management desired, in addition to the incentive set up in the union contract, to publish every three months an honor roll of employees who had produced an exceptionally high volume, in order to add to the monetary benefits a pride incentive.

This was accomplished by means of a control chart on production. Each worker's per cent of production above standard was calculated each month, and an average of three months' figures was calculated and plotted as on Chart XIII. The average chart shows which workers were significantly superior in total output; the range chart shows which ones were *consistent* in their output (a low point on the range chart indicates consistency). The original data was set up in the form shown in Table 18.

TABLE 18

AVERAGE PRODUCTION ABOVE STANDARD

Clock No.	Jan	Feb	Mar.	3-Month Average	3-Month Range
501	2.1%	2.2%	4.2%	2.8%	2.1%
502	2.0	11.6	17.0	10.2	15.0
503	7.9	8.8	4.0	6.9	4.8
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

$\bar{\bar{X}} = 11.2\%$

$\bar{R} = 11.6\%$

$A_2 = 1.023$
$D_3 = 0$
$D_4 = 2.574$

Limits: $\bar{\bar{X}} \pm A_2\bar{R} = 23.0\%, 0\%$

$\bar{\bar{X}} \pm A_2\bar{R} \times 2/3 = 19.1\%, 3.3\%$

$D_3\bar{R} = 0\%$

$D_4\bar{R} = 29.7\%$

Each employee worked on many different short-run jobs during any month, so that the subgrouping method chosen included all variations due to normal differences between jobs; the long period of time covered

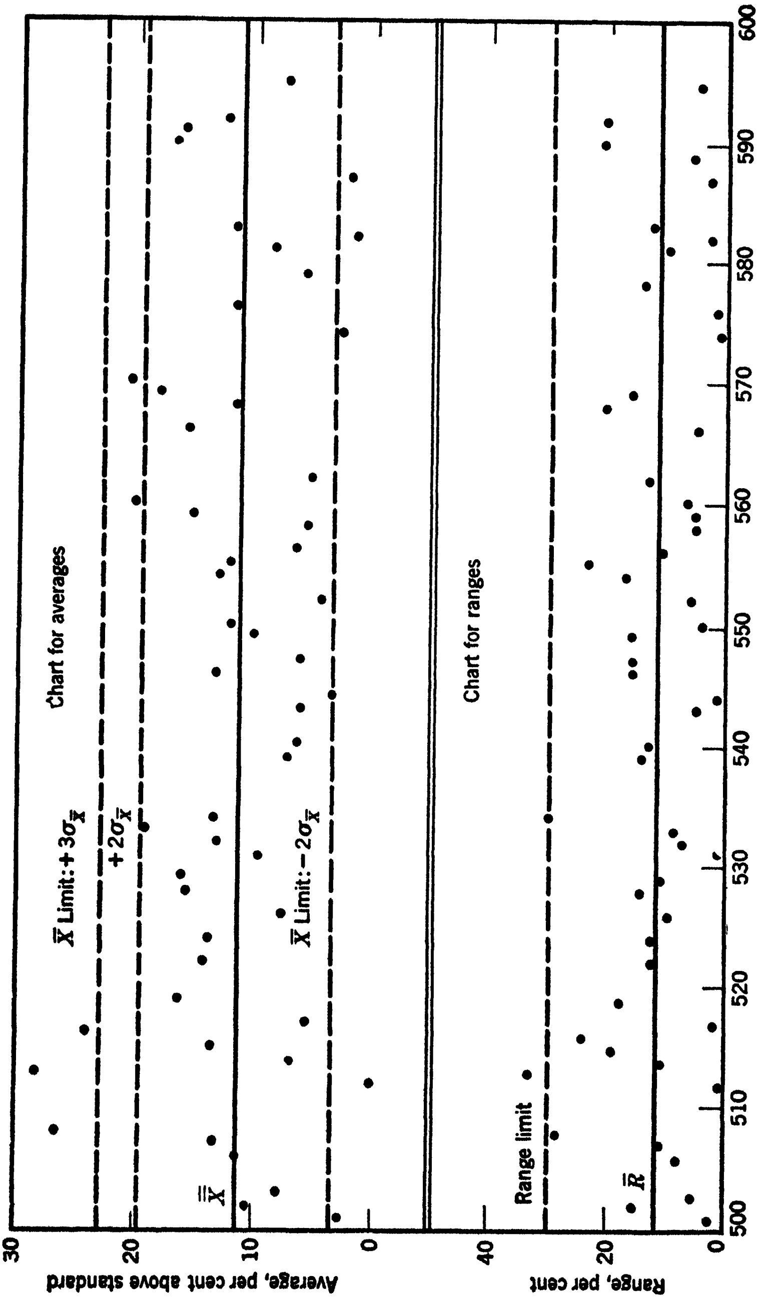


Chart XIII. Production volume of individual workers. Per cent above standard. Average of three consecutive months

(three months) made it possible to include in the sample the normal variations in production caused by each worker's emotional habits and personal circumstances. Excluded from the samples was any significant difference between the workers as to skill, ability, hard work, and other factors determining their volume of output. Such factors, if present in significant degree, would show up as an out-of-control point on the average chart. An out-of-control range point would indicate some instability in the particular worker concerned.

The $\bar{\bar{X}} + A_2\bar{R}$ limit showed three workers with exceptionally large volume: clock numbers 508, 513, and 516. Number 508 was also out on the range chart. The fact that all these points were among the lower clock numbers raised the possibility of long service as an assignable cause. Investigation proved that the three who were out of control were indeed among the oldest employees, and were all recognized as outstandingly skillful. This finding, incidentally, went far toward offsetting the complaint sometimes heard that certain workers were always given the "soft jobs." If employees known to lack skill, training, or hard work had appeared out of control, those complaints would have been justified. Such, however, was not the case. Certainly, acknowledged skill and experience were sufficient reasons for a superior volume of production, without adding the accusation of favoritism.

Number 508's out-of-control range point was found upon investigation to be caused by his working for two weeks upon a special job at standard rates without incentive pay.

For the purpose of preparing an honor roll it was considered advisable to include more than three out of 55 employees. $\bar{\bar{X}} \pm \frac{2}{3}A_2\bar{R}$ limits therefore were used. Five names appeared on the roll, corresponding to the five points outside this closer limit: Nos. 508, 513, 516, 560, and 570. On the other hand, five workers were unusually poor producers, falling below the $\bar{\bar{X}} - \frac{2}{3}A_2\bar{R}$ limit. These were given special attention in order to improve their volume.

This method of separating the *significantly* good and *significantly* poor workers from the mass is especially valuable in building morale, because it enables honor and rebukes to be given on the basis of genuine known reasons and not on the basis of foremen's whims or occasional flashy performance. It puts reward and punishment on a foundation of reasonable concrete causes and appeals to everyone's sense of justice.

CASE HISTORY XIV. FLOW OF PRODUCTION

An application of control charts to production control resulted in a major manufacturing economy in one plant. Costs in the shipping

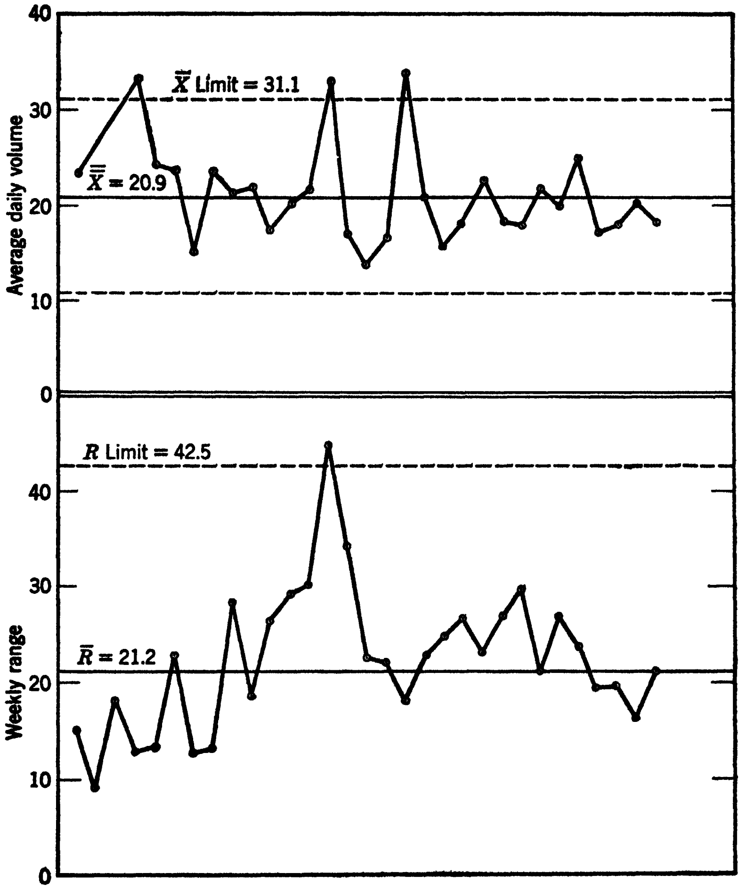


Chart XIV. Flow of production. Receipts into warehouse. Daily average per week (n = 6 days)

department seemed unduly high. The head of the department claimed that irregular deliveries from the factory required him to maintain a larger working force than he needed and to pay a large overtime premium in order to take care of peak deliveries and make prompt shipments. As proof of his contention he submitted a control chart similar to Chart XIV.

Each point on the chart represented the average daily receipts from the factory for one week. Significant differences between weeks showed up as out-of-control points on the $\bar{X}$ chart. Excessive variation from day to day within any one week caused lack of control on the range chart. Both charts were out of control, supporting the shipping department's statement that deliveries from the factory were unreasonably irregular. On the average chart the weeks ending February 28, May 9, and June 6 and on the range chart the week ending May 9 were excessively high. Comparison of payroll records for these weeks with other more normal weeks confirmed the fact of excessive overtime in the shipping department.

Tracing this high-cost factor back into the plant revealed that the assembly department showed unusually high overtime premiums paid in the last week of almost every month.[2] Further investigation revealed a progressive breakdown of the weekly production schedule as the product approached completion. The final assembly department was under continual pressure from factory executives to make as good a showing as possible each *month*, because the accounting records were kept on a monthly basis, and top management gauged performance from monthly rather than weekly records.

When the expensiveness of this attitude was pointed out, and greater emphasis was put on maintaining the *weekly* production schedule, a gradual and consistent improvement in the regularity of factory deliveries took place.

CASE HISTORY XV. COST TOLERANCES

A company which manufactured several thousand different products instituted a cost study to discover the effect of changing conditions on its costs. On each product the following data were gathered: 1939–40 average cost; last year's average cost; and the current year's average cost by quarters. Each of these figures was reduced to a common denominator and made comparable with all the others by expressing it as a per cent of the selling price. The tabulation of these figures was an extensive one; in order to visualize the facts they were put in the form of Chart XV. The tops of the solid lines represent 1939–40 average percentages; the tops of the dotted lines represent last year's averages; the individual dots represent quarterly averages for the current year. The 1939–40 average, the current average, and the current

[2] For a similar study of office overtime costs see the author's paper "Quality Control Applied to Business Administration," *Journal of the American Statistical Association*, June 1943.

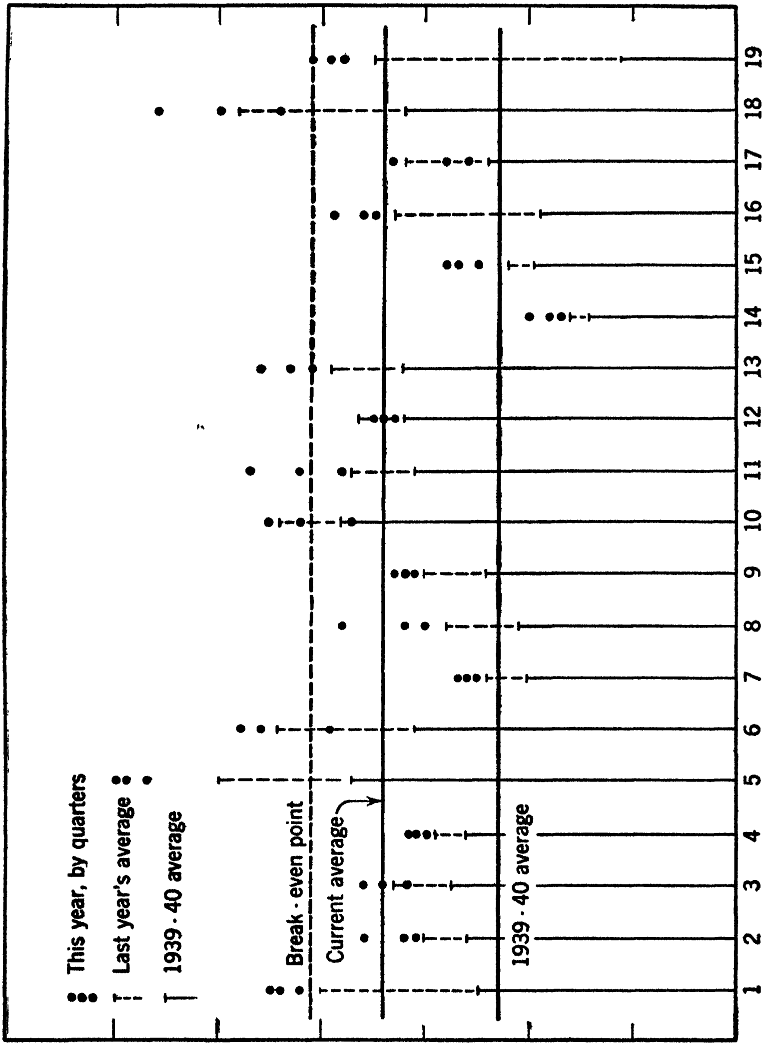


Chart XV. Cost tolerances. Unit costs of products as per cent of selling price

"break-even" point were plotted as shown on the chart. If the average cost were to rise to the "break-even point," all operating profit would be wiped out.

Though this chart is not a control chart in the technical sense, it is a good illustration of the value of graphic statistics. It applies the concept of tolerance limits to costs in a vivid way. Any individual product falling above the break-even point or upper tolerance limit was unprofitable for one or both of two reasons: (*a*) The selling price was too low; (*b*) the factory cost was too high.

Trade practices prohibited price adjustments; rising wages and scarcities of raw materials accounted for many maladjustments. The chart did, however, set up basic facts for long-range planning, and enabled management to concentrate upon the critical products such action as they were able to take under existing conditions. Kept up quarterly, it brought immediately to executive attention any costs that showed signs of getting out of line. It led to an intensive cost-cutting program and setting up a method for gauging the progress toward lower costs.

CHAPTER 6

ORGANIZATION OF A STATISTICAL-QUALITY-CONTROL PROGRAM

CONSULTANTS

Statistics can be put to work in the factory in any of several different ways. One or more specific problems can be tackled, leading by gradual extension to plant-wide applications of statistical techniques; or top management may decide to install statistical controls simultaneously in a co-ordinated program throughout the factory. Either of these two approaches can be carried out by employing consultants from outside the organization, or by utilizing people already employed, or by hiring new employees for the purpose.

Each method has its advantages and its disadvantages. From the management standpoint, the easiest and quickest way to gain the full benefit of statistical techniques is to employ a fully qualified consulting firm to make a complete installation. Such a program, including a thorough educational campaign among executives, foremen, workers, and inspectors, and followed by a sustaining contract for continued advisory service, should be the most efficient way of setting up a complete statistical-control system in the factory. Given full backing, with adequate authority, the consultant who can provide genuine experts in the fields of inspection methods, tooling, process engineering, industrial statistics, sampling plans, and factory management can do a thorough professional job in the shortest time with maximum benefits to the company.

Practical difficulties, however, make this approach difficult. There are few consulting firms in existence; even if there were more of them, their qualifications would need to be given careful consideration, because there are not enough men available who combine with a knowledge of their own special fields the necessary training and experience in the application of statistical methods. Unqualified personnel from outside the organization, who embarked upon an all-embracing program, posing as experts, easily could destroy the effectiveness of the quality-control program permanently. Management in that case would pay a high fee with no advantage gained, at the cost of serious disruption and conflict within the company.

A consultant employed to help in solving a problem of a specifically statistical nature—setting up a sampling inspection plan, for instance—should be very effective. He comes in without threatening plant-wide upheavals, without claiming expertness except in a highly specialized field, without presuming to tell old-time employees how to do their work. Because such a consultant does claim to be a specialist, he really can be one, not only in the theoretical but in the practical aspects of his profession. If he is setting up a scientific sampling inspection plan, he should know by first-hand experience the physical meaning of "randomness" or "representativeness" in a sample, and some of the mechanical difficulties met in getting the right kind of a sample from various kinds of processes. It is easy for the mathematical theorist to assume "randomness," but it is often not at all easy to get randomness in a panful of small parts produced in a machine shop.

If a consulting statistician with practical experience and common sense can be found, the management of a company wishing to install statistical controls can let one job lead to another until a complete system or program has developed with a minimum of misunderstanding among regular employees. Such a "program" could be planned in advance with enough flexibility to meet factory problems as they arise, but the "plan" should appear to be merely one job after another, not a program in the usually accepted meaning.

This somewhat circuitous route to statistical control is recommended, because it can happen that, no matter how thoroughly top management is determined to use the most modern tools in their plant, some opposition will arise from those in lower levels who do not understand and may be incapable of understanding what statistics is and what it can do. The apparently piecemeal approach, if skillfully used, will lull the suspicions and jealousies that otherwise might be aroused by a publicly announced comprehensive "system."

EMPLOYEES

Many firms hesitate to take the risks that are involved in selecting a consultant. They prefer to hire a qualified man to do the job or to train someone already in their ranks.

In certain already established professional fields such as accounting or engineering, the employment of a new man with recognized qualifications to do a specialized job will meet with no opposition and may even be looked upon with favor. If, however, the professional field is not yet well established nor the profession well-defined, and if the so-called professional man seemingly has to interfere with the thinking

and working habits of people who are set in their ways and consider themselves professionals, his path is apt to be a thorny one. The industrial statistician especially employed for carrying out statistical studies in the factory may find himself in just such a predicament. He may find that, although he has the full backing of top management, he does not have the prestige of a consultant, while his responsibilities are heavier. Only if he is a widely recognized authority will he have the prestige necessary for doing fully effective work. Too few men in the United States today have such recognition as industrial statisticians.

One recourse is left to management in seeking to use statistics for solving factory problems: Train a man already employed. He should have a thorough foundation in theoretical statistics, one or more full courses in the specific technique of control charts, a knowledge of the manufacturing process, a scientific and open mind with engineering training if possible, and a position and personality which will make him acceptable to production men, inspectors, and top management. It would be wise to give him a trial period on specific quality and cost problems before setting up under his guidance a full-fledged statistical-quality-control program.

If we assume that it is the avowed intention of management to use statistical methods for gathering and analyzing production data on a plant-wide basis, the remarks that follow outline a few of the principles that, in the author's experience, underly a successful and permanent program.

ATTITUDE

To install and carry on a successful statistical-quality-control program is not an easy task. Unless the attitude with which all the employees of the quality-control department approach their work is right, unless they show clearly and sincerely their wish to do only what is helpful to each person they contact, unless they enable other employees to do their jobs better, neither the best statistics in the world, the most expertly drawn up forms, good organization, influence in high places, executive ability, nor any other asset will make the program a success. Criticism of others cannot be avoided, for quality control is in the unenviable position of a man who advises another man how to do his job. Very often it may seem that the company's interests will be served best and efficiency increased if the bald bad facts are laid on the line and a "non-co-operative" man is bludgeoned by appeal to higher authority. But when the quality-control engineer meets opposition in his efforts to push forward his program, he always should remember that a missionary or a pioneer is judged more by his

actions than his words, that few worth-while causes are won except by persuasion. If he succumbs to the temptation to use force of facts to compel co-operation from a recalcitrant opponent, he will make enemies instead of friends, no matter how right he is. Quality control of the statistical kind can be carried out only by friends who are completely sold on the program and on the quality-control engineer personally.

The members of the quality-control staff are not the only ones who believe they are right. Almost every man in a responsible position is sure that he knows how to do his job. He is apt to resent unasked-for advice, unless it is given with the air of helping rather than criticizing. The quality-control engineer who approaches his job humbly, admitting that he is not an expert on everything, but insisting, tactfully and with determination, that what he does know will be helpful to his fellow workers, will do much to make the program a success.

QUALIFICATIONS

A friendly helpful attitude is an essential ingredient of success, but not the only one. The quality-control engineer, whether newly employed for the job or trained in the plant, must be a good executive, able to inspire his employees and direct their work efficiently. He should know enough about machinery, tooling, inspection, and engineering to discuss with factory people their problems. He must know the technique and philosophy of statistical quality control. He should have a well-grounded knowledge of general statistics and some experience in production or design engineering. If all these qualifications are not found in one man, several men can contribute to the composite knowledge required; but, if it is necessary to use more than one man in supplying all the qualifications for running a successful quality-control program, they must be able to work together as smoothly as a championship football team; and the member of the team who heads the organization must be acknowledged willingly by all to be their captain. Factions within the staff, conflicts of authority, and failure to present a united front easily can prove disastrous.

When the top management of a company, after deciding to give statistical quality control a trial, have found a man or men with the attitudes, knowledge, and practical experience necessary for successful development of the program, the start should be experimental. Beginning in a modest way with one department, either the first department in the order of the manufacturing process, or with the one that seems to need quality control the most, or with final inspection where outgoing quality is determined, the program can be expanded step by step

until the whole plant is covered. Parallel with the development of the program should be the growth of the quality-control staff. At first employing as few people as possible, the staff should be added to only when more help is imperative. The quality-control engineer can do a great service to himself and to his cause by keeping overhead down to the minimum, not only in personnel but also in forms, chart work, filing, routine, and responsibility.

In the matter of responsibility there is an ever-present temptation to assume some of the inspection and supervisory duties in the plant. These temptations should be consciously and deliberately resisted. It is wise to insist that other departments do their own work and carry their own overhead, at the same time proving that every bit of extra work required of other departments by quality control will pay its own way in greater efficiency and better work.

There will arise many occasions when statistical methods can be used in engineering and methods work. Tolerances, for instance, can be set more intelligently and profitably if statistical techniques are called upon to determine what the factory is able to do in holding down variations in the production process. Methods study, which is basically the attempt to improve processes as to quality or quantity of production, needs statistics in collecting and analyzing the facts, and in determining whether or not an apparent improvement is really a significant one. In such projects, however, the function of the quality-control department should be advisory and analytical. The work of gathering and compiling the data should be done by the departments concerned, in the form and manner requested by the quality-control department.

POSITION AND AUTHORITY

How to get adequate authority for carrying on his work is one of the most difficult problems that can face a pioneer—for, in most plants where statistical quality control is being tried for the first time, pioneering is necessary. Most big executives are practical men who insist on having plenty of evidence before they make a decision; and that evidence must be at first hand in their own plant, not derived from some other company's experience. Consequently, the quality-control pioneer must prove his worth before getting authority. Because quality control often demands forms and procedures from the inspection, production, and engineering departments that they are not used to and that seem to demand extra work, and because acceptance of statistical techniques requires a different outlook on manufacturing problems and is apt to disrupt tried and true rule-of-thumb methods,

the officer in charge of manufacturing may hesitate at first to give the quality-control engineer the authority that he deems necessary for carrying on the program.

The quality-control department is in the difficult position of depending for its data upon the inspection department, of which it requires extra work in the filling out of special forms; it depends for its success upon the co-operation of production departments in making special investigations and carrying out its recommendations; it frequently has to persuade the engineering department to change its mind about designs and specifications, upon a type of evidence and logic that is not necessarily self-evident to those untrained in statistics. In order to surmount these obstacles and to make his work effective, the quality-control engineer needs a position and authority in the company that will enable him to meet other department heads as an equal. At the start it is unlikely that he will be granted such authority for a new program, no matter how conclusively it may have proved its value elsewhere. The quality-control engineer who is pioneering in his company usually must accept the fact that he and his work are on trial, and should adapt his demands accordingly.

It matters little where the program starts, whether in inspection, production, or engineering, just so it gets started somewhere. If the quality-control engineer does a good job—and if it is good it is likely to be spectacular—he soon will find himself the equal of other department heads, in fact if not in name. Suppose that the quality-control engineer reports, at the beginning, to the chief inspector. When he finds out-of-control points on his charts, and by eliminating them is able to effect improvements in the manufacturing process that production men realize are practical, it shortly becomes clear that the inspection department is too restricted a sphere for him: he needs direct access to the works manager or general superintendent without having to go through the chief inspector; otherwise the long-established and accepted lines of authority between production and inspection may be violated. If the facts revealed by the control charts are handled tactfully and correctly, the works manager will sooner or later realize their value and will want a more direct organizational approach to the quality-control engineer than through the chief inspector. Then will come the opportunity for quality control to gain its rightful place in the business structure as an equal with other major departments. To work the program into a position of real and equal authority should be one of the chief concerns of the quality-control engineer. When that position has been won, *de facto*, the rest is easy, for then the officer in charge of manufacturing needs merely to accept a *status quo*, an

accomplished fact, in formalizing the authority which the program deserves.

PUBLICITY AND ROUTINE

The quality-control engineer's selling job is a simple one: to present the stories told by his control charts so effectively that action will be taken on them by other departments with confidence and enthusiasm. It does no good, and a great deal of harm, to keep the valuable information gained from quality-control work bottled up in the office, or sent through a single channel. It should be broadcast—judiciously, of course—to as many people as possible, the higher up the better. The author knows of one program that died because no one in authority realized the value of the work being done: the quality-control engineer did not understand the need for publicity, and since he had failed to tell other people about his work he was unable to convince the "big boss" when he was asked to justify his budget.

One of the best ways to get recognition is for the quality-control department to carry on as little routine as possible, doing only what is necessary. Routine forms and reports are to a certain extent unavoidable, but a too elaborate system of records is not to be desired. A few basic charts and tabulations, supplemented by special studies when required, will keep the staff small and the budget reasonable. The same principle applies to work asked of others: the inspection and production departments will like quality control better if they are not asked to take on a great mass of extra record keeping that does not bring obvious benefits.

No fixed rule can be given for the forms and tabulations needed, because they will be different for each plant where a program is being carried on. For small and medium-sized plants the author has found that one form for inspection by attributes and one for variables may be sufficient in process inspection, with different headings, perhaps, for different departments; and another form for acceptance and/or sampling inspection at receiving and final inspection stations. All compiling of inspection reports can well be done in the quality-control office, where tabulations can be set up on the same sheets the control charts are on. If charts are kept at machines, it may be advisable to set them up specially, with particular attention to vividness, readability, and ease of plotting; such charts are usually not satisfactory as permanent office records. Only charts that are really necessary should be maintained. When an operation on which a chart has been kept has reached a state of reasonable statistical control, the chart often can be dropped until obvious trouble again appears. The time

saved can be put to better use in starting a chart on some other operation which needs attention. It is preferable to let a few out-of-control points slip by without being charted than to keep dozens or hundreds of charts that do not tell a story. The original data, however, should be kept in complete tabular form where, for analysis purposes, an experienced quality-control man often can detect assignable causes without the visual aid of a chart.

Reports going out from the quality-control department should be as nonroutine as possible. If, for instance, daily or weekly "hot-spot" reports are made on routine forms, each should be accompanied by a brief explanation telling the story which is implicit in the report. If no "story" is present it may be better not to send out the report, although it may be a real out-of-control spot; this is particularly true if the trouble is one that has occurred recently and is being worked on at the time. Most production men, once they have gotten used to the idea, will welcome facts that help them in their work, but if they already know the facts and are honestly trying to eliminate the trouble, they are apt to resent being told about it again. Or, if the difficulty is known but is apparently insoluble—as when a certain product has to be made on a machine for which it is not adapted, and no other machine is available—then a report pointing out that problem and perhaps insisting on investigation and action will be resented. If, again, manufacturing processes are so badly out of control that dozens of reports have to go out each day, it probably will be good policy to send out only as many of the worst ones as the production supervisors can handle. In the quality-control office, too, the problem of follow-ups may arise. When so many reports go out that they cannot be followed up closely and personally by the quality-control staff they will lose their effectiveness.

Weekly or monthly reports of findings, work done, and conclusions make follow-ups easier. Such reports can well be set up in a standard form and prepared in a routine manner, but they should never be presented in that way, either to an immediate superior or to top management. A special memorandum accompanying the tabulation and describing the high points revealed in the report always will help to get better reception. The memorandum should be brief and clear and should tell an interesting, significant story. It should without exception follow accepted lines of authority. If, for instance, the quality-control engineer is independent, the report and its accompanying note should be addressed to the appropriate officer of the company, with copies to the chief inspector, the works manager, and any others of equal rank who are concerned. In general, reports and memoranda

concerning particular workmen or foremen should be addressed to the department head who has supervisory responsibility in the case, and not direct to the subordinate. In every way avoid action that might be construed as being "out of order" in the organizational setup of the company. Sometimes the real responsibilities do not follow the organization chart, in which case an intimate and practical acquaintance with the various upper strata of employees and their relationships to each other is necessary. This is basically a matter of knowing the top men in the company and their jobs, so that no false step will be taken which might hurt the pride or feelings of an important man, or of one who thinks that because of his position he should be considered important.

A memorandum written specially to and for the man whose responsibility it is to act on it flatters him and therefore exerts a powerful psychological effect. Routine reports may become so familiar as to make little impression, but a personal note usually will be read, and, if it comes from a man who has something worth while to say, it generally will be acknowledged. There is no better way of getting action than to talk to a man personally about one of his problems, particularly if a constructive suggestion can be made toward its solution. Carbon copies of such notes build up an interesting file that can be used, upon occasion, to impress top management with the value of the work being done. Any report that does not immediately arouse the impulse to action probably will be forgotten; and if it is forgotten nothing will be done about it. The quality-control engineer who wants to put his program across quickly and effectively should cut routine to the bone and concentrate on personalized *action* reports. Then the program will not have to be sold. It will sell itself.

RESULTS

Sometimes difficult and puzzling is the problem of how to bring the quality-control work to the attention of top management in an impressive way. That this is necessary goes without saying, but, if the program is under the wing of inspection or production or engineering or some other manufacturing function, it is not always easy to know just how to break through the lines of authority and get a hearing at the very top. Even if the quality-control engineer is independent, with direct access to the man in charge of manufacturing, or to someone of similar rank, it is necessary to approach him in the right way.

First of all, reports submitted to top management, through whatever channel, should be practical, couched in language that executives

understand, nontechnical, and factual. The "pep-talk" type of report should be avoided. Extravagance or overenthusiasm have no place in a project which claims to be scientific; one too optimistic statement can discredit the whole quality-control program in the mind of an important man, because, if he knows anything at all about it, he knows that it is based on statistics and mathematics. He therefore expects from it more accuracy and better judgment than he does from other departments. For the same reason he expects more results and more value for the money spent than from other departments. The author has discovered that executives generally are inclined to be critical of statistical quality control unless it performs near-miracles. This demanding of extra value is partly due to the technical nature of the work, which to most executives is new and somewhat strange, and partly to the sometimes exaggerated claims that are made for it by its converts. Quality control is not a religion, nor a cure-all, nor infallible; and, while those who believe in it do not say it will solve all problems, that impression sometimes is left by those who let their enthusiasm run away with their good judgment. Plain factual statements of what quality control can do and has done, based upon sound facts, correct statistical analysis of them, and good judgment about the material and human possibilities in the plant, should get the program off to a good start; and similar reports from time to time telling about what the program is doing should keep it in the limelight and win strong, increasing support from top management.

Secondly, these progress reports are usually most effective if the results are evaluated in terms of dollars and cents. An executive will be interested to know that the amount of bad work in a certain department has dropped off 50 per cent, but he probably will be more impressed if it is pointed out that the company is thereby saving $10,000 a month in increased production and $5,000 in reduced scrap, a total of $15,000 a month. Such a statement often can be made more striking if the saving is compared with the cost of the quality-control program in that department.

Thirdly, the quality-control engineer, in order to get his message to top management, should enlist support from as many different directions as possible. One of these, often neglected in quality-control work, is the sales department. The outgoing quality of a factory's product is of vital concern to the sales manager, and his co-operation and support should be solicited and encouraged by the quality-control engineer. In the matter of setting acceptable economic standards, both as to over-all allowable per cent defective and as to specific defects that make the product unsalable or unusable, the sales manager

should be consulted. No quality committee is complete without his presence and active participation. Through him it is often possible to reach the ear of top management from a new direction, one that may offset unfavorable reactions from some hurt or prejudiced member of the factory staff. Here, indeed, is one of the most constructive jobs that statistical quality control can do in any company: to provide an avenue by which sales and manufacturing and inspection can approach each other more closely in their objectives, become more understanding of each other, and thus synthesize their viewpoints into a unified policy. The emphasis that quality control places upon the *economic* aspect of quality—the compromise between what it pays to do and what it does not pay to do—effectively concentrates the aims of both sales and production departments on the real purpose of business, which is to make a profit. If the quality-control engineer accomplishes nothing more than this, namely the drawing together of the sales manager, the works manager, and the chief inspector to a meeting of minds on what the company can and should do, it will have proved its value and paid its way many times over.

THE PERSONAL APPROACH

The personal approach to quality problems encourages co-operation between department heads, sells the program to the men at the top of the executive pyramid, and with tact and patience smooths the path toward full and effective use of statistical-quality-control methods. But it goes much deeper than that. It reaches down to the primary causes of bad work. Even the most automatic operations in the factory involve some of the human element in setups, maintenance, and inspection; and as the operations become less and less automatic the human element increases. People, as workers, are ultimately responsible in some way for all bad work. It is logical, therefore, that the analytical part of quality control, the determination of causes of variability, should emphasize strongly the study of operator or workman characteristics. In the author's experience this has been a more fruitful approach than machine behavior or product quality or raw material defects.

In the literature of quality control to date this line of attack on the problem of bad work has largely been ignored. Several books and many articles have been written covering thoroughly the statistical aspects of quality control, the application to machines, products, raw materials, and the like, but very little has been published on the value of the method in analyzing the work behavior of individual factory

employees, both productive and supervisory. It is poor work behavior —carelessness, laziness, lack of confidence, fear, poor instruction, poor supervision, physical or mental maladjustments, inexperience—which causes much bad work; and it is the reverse of these which produce high quality. The type and condition of machinery, the nature of raw materials, satisfaction with wages and working conditions, and all the elements that go to make up morale have a strong bearing on the quality of work. But it is surprising how, under difficult and unpleasant conditions, a good workman can turn out superior work if he is given the help and encouragement he needs.

In one quality-control program, this personal approach was used very effectively. In the department where the program was being given a trial, there were 20 men working. The first step that the quality-control engineer took was to set up an individual control chart for each man. When he had accumulated data for three months he wrote a four-page memorandum to the department head pointing out, on the basis of his chart records, the characteristics of each of the 20 men. Five of them (he gave names and facts to support his statement) he found were excellent workmen, needing little or no supervision. Eight of them were good, of whom six were fairly consistent, but two (again giving names) did superior work most of the time with occasional flare-ups of poor work; in one of the two cases the flare-ups came when the man was taken off his regular machine and put on another one; the other erratic operator had trouble only with certain jobs. The remaining seven workmen were poor, each showing a specific cause for his bad work. One, for instance, had trouble whenever he started a job, but after he settled down to it turned out a fair quality; another was a new employee with previous experience elsewhere who was showing steady improvement; a third was about to be drafted; a fourth showed lack of training; a fifth was a helper recently promoted to full operator status; and so on right down the list.

The facts that made this report so impressive and valuable to the department head were uncovered by clues obtained from the control charts; the "engineering" investigation (in this case along personnel lines) showed up clearly in each case what the assignable cause was. When the cause had become known, appropriate action was taken to correct it. This is, of course, exactly what statistical quality control can and should do. Applied to people, it works as well as it does on machines. Within a month from the time the report came out the amount of bad work being done in the department dropped from around 6 per cent to less than 2 per cent, which, for the particular process, was extraordinarily good.

This reduction of bad work to one-third its former level was achieved through improved supervision. Foremen were instructed to give the seven poor workers three fourths of their time and to spend most of the other fourth on the good workers, leaving the superior men with very little supervision. Each worker who needed watching was guided by the foreman in a way which would offset his particular weakness. Two or three incorrigibles were given releases. The foremen, too, were put on their mettle; one, who was incompetent, was laid off. Supervision was improved, and its cost was reduced while better work than ever before was turned out by the direct workers.

In another plant, the quality-control program almost failed because in the setting of inspection standards the reactions of the workers involved had not been considered. Here, as in many cases, one of the first jobs that had to be done before quality control could really get started was to set up objective standards for inspection during the process. This particular operation was almost entirely manual in nature, depending largely upon the care and skill of the individual workers. The finishing work performed in this department had to pass final inspection, where standards were supposed to be very strict, although none of the bad work was critical in the usual sense.

The quality-control engineer, in consultation with the head of the final inspection department, set up three classifications, *A*, *B*, and *C* for finishing defects. Since process inspection in the finishing department had been more or less nominal, it soon developed that all the work done fell into classes *B* and *C*, with most of it in class *C*. As a result, about three fourths of the work began coming back to the finish department from final inspection for reworking. The added load on the workers brought on a revolt, and a strike was narrowly averted.

It was an embarrassing moment to the quality-control engineer when he was asked for an explanation of this incident by the vice-president in charge of manufacturing. When he asked for and was given another chance he went about it more tactfully. He personally spent a week collecting samples of typically defective work produced by the men, and when he had accumulated about a hundred samples he asked the vice-president, the chief inspector, and the department head in for a conference on the subject of standards. The four of them sat down for several hours examining and discussing the faults exhibited in the samples. Next, the department head called a meeting of his foremen and got their agreement on the standards to be used. The agreed-upon defective samples next were submitted to the sales manager for his approval. Then the foremen were given photographs of the selected samples to show their men, and another set of prints

was given to the inspectors for their guidance. When all the *people* involved had agreed as to what could be done economically to produce an acceptable finish, the standards were put into effect. This time there was no strike; in fact the quality of the work done improved markedly, and the morale of the department rose to a new high.

DOUBLE-CHECKING ON THE FACTS

No matter how tactfully people are handled, nor how smoothly the program seems to be going, the quality-control engineer must be sure that behind every move he makes, behind every report he issues, there are solid facts and sound analysis. He cannot afford to be tripped up on his statements: one such error will offset nine correct statements. In order to get solid facts it is necessary to check, cross-check and double-check the work, to investigate thoroughly and reason logically before putting any conclusion on paper. In order to get sound analysis, it is necessary to use sound statistical techniques. For these two requirements there is no substitute.

In checking his work, the author has found it desirable not to take unsupported evidence at its face value; by unsupported evidence is meant data coming from a source that cannot be statistically confirmed. If inspection at the hammers is set up in the forge department, the inspection records should be checked with those at a subsequent operation such as trimming; in this case the trimming inspectors should look not only for trimming defects but for forging defects as well. Every job that is forged can be compared with the quality shown at trimming, the comparison being made by the use of control charts on which both inspection results are shown, the trimming reports being plotted on the forging control charts in a different color. Both the forging and the trimming points must be inside the control limits before either inspection is accepted as valid.

Sometimes it is not the process but the inspector that needs to be checked. If a certain product has been running well above the zero line on a p chart, and suddenly a point appears below the lower control limit, it is unlikely that the cause is improvement in the process. More likely, it will be a flaw in inspection. Or, if one inspector is suspected of failing to follow instructions, his reports can be checked with those of another inspector on the same operator, job, and machine; or the quality-control engineer personally can sample-inspect the lot after the inspector has inspected it.

Another place where statistical checking is necessary is at final inspection. If certain defects are really critical and must be 100-per-cent-

inspected, a sample should be taken before and after the 100 per cent inspection. The preinspection sample can be used as a guide for determining how many defects should be found at detailing; the postinspection sample gives a check on the effectiveness of the detailing.

In accepting the results of engineering investigations of trouble spots, the quality-control engineer should use caution. He should confirm engineering conclusions by getting information from other sources, unless the conclusion is obviously correct. One promising quality-control program ran into difficulty because the wrong department was blamed for the trouble. This particular fault was off-center broaching. An investigation disclosed that apparently the jig was causing the bad work, whereupon the tooling department was called upon to provide a satisfactory fixture for the broaching machine. A new one was made and tried, but the bad work persisted. A further study showed that carelessness at a previous operation was the true cause of the trouble; but by that time the quality-control engineer's report had become public property and, though he was not to blame, he had to shoulder the responsibility. In another instance an out-of-control point was apparently due to the fact that the foreman on the job had not checked the work with a gauge but had depended on visual inspection; he claimed that no gauge was available. The quality-control engineer checked up on the gauges and found that actually three or four were available, that probably the foreman simply had not asked for one. Unfortunately it turned out that, shortly before this incident occurred, there had been a change in specifications; that the gauges that had been available had been old ones and that, in fact, a gauge for the new specifications was not available. However, the damage had been done when the report went out to the department head that his foreman had neglected to get a gauge when one was to be had. After several such instances had occurred, the quality-control engineer lost the confidence of the factory men; his program never recovered from the loss of prestige.

Mistakes of this kind sometimes can be excused if they are known only to department heads and if the quality-control program honestly is trying to do a good job, the errors arising because of poor judgment rather than poor intentions. If, however, the top executive hears about them from an injured party, or if a mistake is found in a report, the damage may be greater. An enthusiastic memorandum to one executive pointed out that in six months department A had shown a 75 per cent decrease in the amount of bad work, comparison being made between a week just before negotiations for a new union contract were begun and a week after the new contract was signed. This fact

occurred to the executive (though it had not been stated in the report) and he recalled that in the earlier week the plant had been almost closed down owing to friction between workers and management. He questioned the validity of the report and as a result began to doubt the value of the quality-control program.

An even more serious flaw developed in another program. This was at receiving inspection, where parts from various subcontractors were sample-inspected, with the acceptable quality set at 1 per cent defective. Statistical tests were set up, based upon sampling tables, providing that only a certain number of defective pieces were to be permitted in the sample. It happened that, because of the pressure of consumer demands, the production departments were allowed to overrule inspection if in their opinion the parts were needed badly enough. The trouble was that production overruled inspection two thirds of the time, and the quality-control engineer allowed this situation to continue for months without doing anything about it.

A test that functions only one third of the time is not a valid test and should not be used. The quality-control engineer did not realize this fact. He should have known that his whole program was being nullified by the action of the production men and should have taken steps to correct the situation. If he had thought it through, he would have understood that inspection standards in this case were not in line with company policy; it would have been a relatively simple matter to persuade the management to set up realistic rather than idealistic standards. But at the end of nine months it became apparent that the program was not producing results, and, though personal relationships had been well handled by the quality-control engineer, the president of the company decided the project was not worth continuing.

SOUND STATISTICS

It has been said of quality control that it is nine-tenths engineering and one-tenth statistics. Whatever the proportion of engineering may be, the author feels that sound statistics carries more than 10 per cent of the weight in a successful program. Any scientific method that is not on a firm theoretical basis is apt to fail in a crucial test, and this applies to statistical quality control as well as to engineering and every other scientific field. One or two illustrations will clarify this very important principle.

In Company X, a very large corporation, a young executive was trying to get approval for a statistical-quality-control program. He had made a number of special studies, all of which had met with ap-

proval. One day the company received from a certain vendor 500,000 rivets. They passed receiving inspection, but it was discovered too late that they were soft; the problem was how to find out which of 75 supply depots in the plant had received the bad rivets. The young executive offered to set up a sampling scheme that would determine where the faulty rivets were. The scheme was carried out and 26 depots were found to contain bad rivets. These were detailed in order to remove the defective parts. Unfortunately, during the next week three more depots were found to contain some of the rivets, the discovery being made during the manufacturing process. This failure threw doubt on the value of sampling inspection and resulted in a decision not to install a quality-control program.

The mistake was made because the young executive had not asked the question, "What risk is allowable that some bad rivets will not be located by the sampling scheme?" Such a question was necessary, because there is always an error attached to sampling, an error which however, it is possible to estimate and which can be made as small as desired by selecting the correct sample size. The number of pieces taken for inspection from each of the 75 supply depots according to the scheme used was only large enough to hold the error to 10 per cent; that is, nine out of ten of the defective bins would have been discovered. Management, however, was not willing that any of the rivets be undiscovered. The mistake was made because the man who drew up the sampling plan was not thoroughly familiar with the principles of sampling theory, and therefore used a good method in an incorrect way.

A final example of the dangers that a lack of sound statistical training abound in is the story of a recently appointed quality-control engineer who was unable to fulfill the promises he had made, because he used the wrong kind of control chart. His problem was that on each assembled unit of product many different kinds of defects were found, some of which were expensive to correct and others of which were very minor. It was not practical to run a control chart on every type of defect, because too many charts would have been required. He combined the different types of defects into a weighted total by assigning arbitrary weights to each kind of fault. But when he plotted the resulting points on his p chart as a per cent defective, it was found that the limits had no meaning. Out-of-control points, which should have pointed at trouble in the process, often proved to be the result of normal variations; and many points which were in control nevertheless turned out to be real indicators of trouble. If the quality-control engineer himself had discovered these flaws in his charts it might not have been

such a serious matter; but, since they were discovered by the production men as a result of making investigations that showed no trouble, or of not making investigations when trouble existed, the factory lost confidence in the quality-control program. Because the quality-control engineer did not realize his predicament and did not see that his statistics were poor, the program was dropped.

There are several alternative ways in which the problem of charting different kinds of defects on a single unit of product can be handled. Two or three of the most critical or expensive defects can be charted, each one on a separate chart, with either p charts or $\bar{X}$ and R charts, depending on whether inspection is by attributes or by variables, all others being lumped into one chart; another method is to calculate the total cost of rework or scrap on a succession of units and run an $\bar{X}$ and R chart on the total cost; a third way is to classify the defects into categories such as critical, major, and minor, and run a chart on each category. Which of these alternatives is used will depend upon the purpose of the chart; but certainly it is dangerous to try to combine many different systems of causes on one chart, because the chart then will not describe any system of causes at all. In describing the physical characteristics of horses and donkeys it would be misleading to use the appearance of a mule as the criterion.

Enough has been said to emphasize the necessity of a quality-control program being based on sound statistical methods. So numerous are the pitfalls to a man inexperienced in statistical fundamentals that no program should be undertaken unless adequate statistical knowledge is available. Equally necessary are solid well-tested facts. If the man who is running the program is not only factually minded but also meticulous in the use and presentation of his facts, the program will be well on the road to success.

INDEX